

# **Sicheres Fahren durch unsichere Automation?**

Was leistet die Rückmeldung von Automationsunsicherheit bei der Übernahme der Kontrolle  
vom teil-/hochautomatisierten Fahren?

Von der Fakultät für Lebenswissenschaften  
der Technischen Universität Carolo-Wilhelmina zu Braunschweig  
zur Erlangung des Grades  
eines Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Matthias Johannes Heesen  
aus Duisburg

1. Referent: Prof. Dr. Mark Vollrath  
2. Referent: Prof. Dr. Klaus Bengler  
eingereicht am: 14.08.2020  
mündliche Prüfung (Disputation) am: 16.12.2020

Druckjahr 2021

## **Vorveröffentlichungen der Dissertation**

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

## **Publikationen**

Beller, J., Heesen, M., & Vollrath, M. (2013). Improving the driver-automation interaction: an approach using automation uncertainty. *Human factors*, 55(6), 1130–1141.  
<https://doi.org/10.1177/0018720813482327>

## **Tagungsbeiträge**

Heesen, M. & Baumann, M. (2013) Uncertainty feedback as component in the interaction design for highly automated vehicles. In: 55th Teap Vienna. Pabst. Tagung experimentell arbeitender Psychologen, 24. - 27, Wien.

Heesen, M. & Baumann, M. (2012) Rückmeldung von Automations(un)sicherheit als Möglichkeit zur Verbesserung der Kontrollierbarkeit von Automationsfehlern bei hoch- bzw. teilautomatisiertem Fahren. Paper presented at 8th Workshop Fahrerassistenzsysteme 2012 Walting.

Heesen, M, Keich, A. & Baumann, M. (2012) Likelihood alarms in highly automated vehicles: Strategies to prepare the driver for erratic automation behaviour. In: Traffic and Transport behaviour: interaction between theory and practice. International Conference on Transport and Traffic Psychology - ICTTP, 29.-31. Aug. 2012, Groningen, Niederlande.

Heesen, M., Keich, A. & Baumann, M. (2012) Likelihoodalarme im Fahrzeug: Strategien zur Vorbereitung des Fahrers auf Automationsfehlverhalten bei automationsunterstütztem Fahren. In: TeaP 2012, Seite 285. Pabst. Tagung experimentell arbeitender Psychologen, Mannheim, Deutschland.

Heesen, M., Beller, J., & Flemisch, F. (2011). Making automation surprises less surprising: Rückmeldung von Automationsunsicherheit als Vertrauensinformation bei assistiertem und hochautomatisiertem Fahren. *Reflexionen und Visionen der Mensch-Maschine-Interaktion–Aus der Vergangenheit lernen, Zukunft gestalten*, 9, 275-280.

Heesen, M., Beller, J. & Flemisch, F. (2011) Making automation surprises less surprising – Förderung der Erwartungshaltung von Automationsfehlverhalten durch Rückmeldung von ereignisgebundener Automationssicherheit. In: *Reflexionen und Visionen der Mensch-Maschine-Interaktion - Aus der Vergangenheit lernen, Zukunft gestalten*, 33, 32-33. VDI Verlag GmbH Düsseldorf. 9. Berliner Werkstatt Mensch-Maschine-Systeme, Berlin, Deutschland. ISBN 978-3-18-303322-5

# **Inhalt**

Danksagung .....	7
Abkürzungen .....	8
1 Hintergrund und Ziele der Arbeit .....	9
1.1 Begriffsklärung Automationsfehler & falsche Alarme .....	12
1.2 Ziele der Arbeit .....	12
2 Theoretischer Hintergrund .....	13
2.1 Fahrzeugautomation und die Rolle des Fahrers .....	13
2.2 Fahrzeugautomation und Situationsbewusstsein .....	15
2.3 Automationawareness – Erwartung des Automationsverhaltens .....	15
2.4 Situationsbewusstsein & Automationsgrade .....	16
2.5 Falsche Erwartung des Automationsverhaltens - Overtrust & Complacency .....	17
2.5.1 Trust und Complacency bei Fahrzeugautomation .....	19
2.6 Kontrollübernahmezeiten .....	20
2.7 Die Problematik der „rechtzeitigen“ Rückmeldung .....	23
2.7.1 Das Vorhersagedilemma – Signalentdeckungstheorie .....	23
2.7.2 Kontrollrückgabe oder dynamische Übernahmebereitschaft? .....	25
2.8 Alarmzuverlässigkeit und Vertrauen in das Alarmsystem .....	26
2.8.1 Dimensionen des Vertrauens in ein Alarmsystem: Compliance & Reliance .....	27
2.8.2 Der Cry Wolf Effekt .....	28
2.8.3 Untersuchungen zur Wirkung falscher Alarme im Kraftfahrzeug .....	29
2.9 Lerntheoretische Grundlagen des Cry Wolf Effektes .....	30
2.9.1 Die Evidenztheorie des kausalen Lernens .....	32
2.10 Möglichkeiten der Entstehung eines Cry Wolf Effektes zu entgegnen .....	34
2.10.1 Einfluss von Vorwissen auf das Erleben falscher Alarme .....	36
2.10.2 Auswirkung der Alarmgestaltung auf den Cry Wolf Effekt .....	37
2.11 Einfluss von Likelihoodinformation auf die Hypothesenbildung .....	40
2.11.1 Likelihoodinformation durch Rückmeldung von Automationsunsicherheit ....	42
2.11.2 Likelihoodinformation bei assistiertem & hochautomatisierten Fahren .....	45
2.11.3 Rückmeldung einer Begründung für Kontrollrückgaben .....	49
2.12 Zusammenfassung der Theorie .....	51
3 Leithypothesen .....	54
3.1 Auswirkungen auf die situative Erwartung von Automationsfehlern: .....	54
3.2 Auswirkungen auf die generelle Erwartung von Automationsfehlern .....	56
4 Übersicht der Studien und Zuordnung zu Leithypothesen .....	57
4.1 Übersicht Studie 1 .....	58
4.2 Übersicht Studie 2 .....	59
4.3 Übersicht Studie 3: .....	61
5 Basis-Untersuchungsparadigma .....	62
5.1 Methodik zur Untersuchung der Cry Wolf Anfälligkeit .....	62
5.1.1 Variation der Erfahrung „falscher“ Alarme .....	66
5.1.2 Variation der Erfahrung tatsächlicher Automationsfehler .....	67

5.1.3	Messung des Cry Wolf Effektes .....	68
5.2	Methodik zur Untersuchung der Effekte auf das Vertrauen in die Automation .....	69
5.2.1	Trustmessung .....	71
6	Versuchsmaterial & generelles Vorgehen .....	71
6.1	Fahrsimulator .....	71
6.2	Umsetzung der Rückmeldungssituation .....	73
6.3	Fahrerablenkung durch Nebenaufgaben .....	74
6.4	Fragebögen und weitere Messinstrumente .....	76
6.4.1	Trustfragebögen .....	76
6.4.2	Erfassung des situativen Vertrauens .....	77
6.4.3	Erfassung der Akzeptanz in die Automation .....	77
6.4.4	Erfassung der Kooperation .....	78
6.5	Generelles Vorgehen: Stichproben .....	78
6.6	Generelles Vorgehen: Datenauswertung .....	79
6.6.1	Varianzanalyse & eingesetzte Korrekturverfahren .....	79
6.6.2	Non-Parametrische Tests .....	79
7	Studie 1 .....	80
7.1	Ziele Studie 1 .....	80
7.2	Methode Studie 1 .....	83
7.2.1	Versuchsdesign je Teilziel: .....	84
7.2.2	Gesamtversuchsschema .....	88
7.2.3	Vorgehen .....	88
7.3	Ergebnisse Studie 1 .....	90
7.3.1	Stichprobe .....	90
7.3.2	Ergebnisse Kontrollierbarkeit .....	90
7.3.3	Ergebnisse Gesamtvertrauen: .....	95
7.3.4	Ergebnisse: Akzeptanz .....	98
7.4	Diskussion Studie 1 .....	100
8	Studie 2 .....	107
8.1	Ziele Studie 2 .....	107
8.2	Methode Studie 2 .....	111
8.2.1	Versuchsdesign für Ziele 1-5: .....	112
8.2.2	Gesamtversuchsschema / Design .....	113
8.2.3	Abhängige Variablen: .....	114
8.2.4	Vorgehen .....	115
8.3	Ergebnisse Studie 2 .....	116
8.3.1	Stichprobe .....	116
8.3.2	Ergebnisse der Kontrollierbarkeit von Automationsfehlern .....	116
8.3.3	Ergebnisse der Nebenaufgabenleistung .....	119
8.3.4	Ergebnisse situatives Vertrauen .....	125
8.3.5	Ergebnisse Gesamtvertrauen .....	128
8.3.6	Ergebnisse Akzeptanz der Automation .....	129

8.4	Diskussion Studie 2.....	130
9	Studie 3 .....	137
9.1	Ziele Studie 3 .....	137
9.2	Methode Studie 3 .....	142
9.2.1	Kontrolle von Sekundärvarianz .....	142
9.2.2	Versuchsdesign je Teilziel .....	145
9.2.3	Gesamtversuchsschema .....	150
9.2.4	Vorgehen .....	151
9.3	Ergebnisse Studie 3: .....	153
9.3.1	Stichprobe: .....	153
9.3.2	Ergebnisse der Kontrollierbarkeit von Automationsfehlern .....	153
9.3.3	Ergebnisse der Reaktionszeit in der Miss-Situation .....	159
9.3.4	Ergebnisse Nebenaufgabenbearbeitung.....	160
9.3.5	Ergebnisse situatives Vertrauen .....	163
9.3.6	Ergebnisse Gesamtvertrauen in die Automation .....	165
9.4	Diskussion Studie 3.....	168
10	Gesamtdiskussion .....	181
10.1	Auswirkungen auf die situative Erwartung von Automationsfehlern.....	181
10.2	Auswirkungen auf die allgemeine Erwartung von Automationsfehlern .....	187
11	Schlussfolgerung .....	189
12	Ausblick .....	191
13	Referenzen.....	192

## **Danksagung**

Es ist geschafft! Zumindest liegt diese Arbeit nun fertig in meinen Händen. In der langen Zeit, die mich diese Arbeit begleitet hat, haben mich auch viele Menschen begleitet und jeweils auf Ihre ganz eigene Art und Weise dazu beigetragen, dass diese Arbeit fertig werden konnte.

Zu Beginn waren da Dr. Frank Flemisch und meine Kollegen aus der Gruppe Systemergonomie und Design beim DLR in Braunschweig. Frank hat mit seinen Ideen maßgeblich dazu beigetragen, dass diese Arbeit überhaupt ihre Idee fand. Mit Anna Schieben, Johann Kelsch, Marc Wilbrink, Stefan Griesche, Gerald Temme und Tobias Hesse habe ich mich immer wieder rege austauschen können. Durch sie fand ich die ein oder andere Inspiration für diese Arbeit.

Ohne die tatkräftige Unterstützung von Julian Schindler bei der Erstellung der Simulationsszenarien, hätte diese Arbeit nie entstehen können.

Dr. Martin Baumann hat auch einen beträchtlichen Anteil an der Entstehung dieser Arbeit, indem er mir einerseits die Zeit gab die Einzelstudien dieser Arbeit durchzuführen und andererseits bei der Planung und Auswertung stets mit Rat und Tat zur Seite stand.

Bei Prof. Dr. Mark Vollrath möchte ich mich nicht nur dafür bedanken, dass er als Mentor zur Verfügung stand, sondern auch dafür, dass er immer wieder Zeit fand mir wertvolles Feedback zu dieser Arbeit zu geben.

Bei Prof. Dr. Klaus Bengler und Prof. Dr. Frank Eggert möchte ich mich bedanken, dass sie als Referenten zur Verfügung standen bzw. den Vorsitz über die Promotionskommission übernahmen.

Ich möchte mich auch bei Johannes Beller und Alexandra Keich bedanken, die mir, im Rahmen ihrer Abschlussarbeiten, bei der Durchführung der Studien tatkräftig zur Hand gegangen sind.

Nicht nur eine Wohnung und eine gute Zeit in Braunschweig, sondern auch die ein oder andere Diskussion zum Thema Dissertation habe ich mir mit Andro Kleen teilen dürfen.

Schlussendlich und vor allem, möchte ich meiner Frau Marie und meinen Kindern Moritz und Mathilda danken, für ihre unendliche Geduld, die ich ihnen abverlangte, wenn ich mich immer und immer wieder mit dieser Arbeit beschäftigte.

## **Abkürzungen**

ACC = Adaptive Cruise Control

TTC = Time to Collision

THW = Time Headway

ANOVA = Analysis of Variance – Varianzanalyse

Km/h = Kilometer pro Stunde

m/s<sup>2</sup> = Beschleunigung / Verzögerung

s = Sekunde

min = Minute

p = Irrtumswahrscheinlichkeit statistischer Testverfahren

Df = Degrees of Freedom (Freiheitsgrade)

Ego-Fahrzeug = Fahrzeug das die Versuchsperson selbst steuert

HC = Heading Control – Spurhalteunterstützung

Eta = partielles Eta – Effektstärke eines mittels Varianzanalyse identifizierten Effektes



# **1 Hintergrund und Ziele der Arbeit**

Die Automatisierung der Fahraufgabe gilt als eine der größten Entwicklungspotenziale für Automobilhersteller weltweit. Sie kann zu einer effizienteren Nutzung von Verkehrssystemen beitragen z.B. durch Platooning, sie kann eine Verringerung von Unfällen und Unfallfolgen durch Systeme der aktiven Sicherheit ermöglichen und sie kann einen Zugewinn an Zeit für die Fahrer bedeuten, nämlich dann, wenn das Fahrzeug über weite Strecken auf einen Fahrer verzichten kann.

Der Weg zu höherer Fahrzeugautomation birgt aber Hürden. Komplexe, dynamische und äußerst variable Verkehrssituationen stellen eine solche Herausforderung für maschinelle Sensorik, Dateninterpretation und Entscheidung dar, dass ein der Situation angemessenes Verhalten der Automation nicht in allen Situationen in die sie geraten kann, mit ausreichender Sicherheit gewährleistet werden kann (z.B. Christoffersen & Woods, 2002). Selbst nach Millionen von zurückgelegten Testkilometern scheitern Versuchsfahrzeuge noch regelmäßig an diversen Fahrsituationen und Wetterbedingungen (z.B. Gomes, 2014).

Nach aktueller Rechtsauffassung ist ein vollautomatisiertes System nur dann zulässig, wenn es in sämtlichen Fahrsituationen ohne einen Eingriff des Fahrers auskommt. Solange dies noch nicht garantiert werden kann, muss der Fahrer als Sicherheitsebene funktionieren und in der Lage sein, die Kontrolle des Fahrzeuges in einem bestimmten Zeitrahmen zu übernehmen (z.B. Gasser et al. 2012; NHTSA, 2013; SAE, 2014). Dieser Zeitrahmen muss so bemessen sein, dass der Fahrer ein ausreichendes Situationsbewusstsein bezüglich der relevanten Objekte in der Umgebung und deren zukünftigen Verhaltens (z.B. Baumann & Krems, 2005) aufbauen kann.

Der Bildung eines notwendigen Situationsbewusstseins für eine angemessene Kontrollübernahme stehen eine Reihe von negativen Automationseffekten auf das Erleben und Verhalten des Fahrers entgegen. Stichworte in diesem Zusammenhang sind "Pilot (Driver) Out of the Loop" und "Loss of Situational Awareness" (z.B. Endsley & Kiris, 1995, Niedere & Vollrath, 2011, Kassner & Baumann, 2011), Overtrust in Automation (z.B. Lee & See, 2004), Complacency (z.B. Parasuraman, Molloy & Singh 1993).

Google zog z.B. in Erwägung beim „Driverless Car“ ganz auf einen Fahrer und somit auf Lenkrad und Pedale zu verzichten, eben weil Googles bisherige Erfahrung zeigt, dass der Fahrer bei hoher Automation nicht in der Lage ist die Fahraufgabe adäquat zu übernehmen (Markoff, 2014).

Fahrer haben aufgrund der genannten Automationseffekte kein ausreichendes Automationsbewusstsein, das Situationsbewusstsein bezüglich der Automation und

wie diese sich in bestimmten Situationen in der Zukunft verhalten könnte. Sie kennen die genaue Funktionsweise einer Automation weder explizit, z.B. durch adäquate mentale Modelle welche Umgebungseinflüsse ein Problem für die Automation darstellen, noch können sie implizit kaum eine korrekte Erwartung über die Leistungsfähigkeit einer Automation aufbauen. Sie trauen der Automation häufig mehr zu, als sie eigentlich zu leisten im Stande ist. Reagiert eine Automation in einer Situation unangemessen so sind Nutzer der Automation daher häufig überrascht und unvorbereitet. Ein Phänomen das bereits von Sarter, Woods & Billings (1997) als „Automation Surprise“ beschrieben wurde.

Eine Möglichkeit, die Bildung eines angemessenen Automationsbewusstseins zu unterstützen, sind Rückmeldungen bezüglich der eigenen Leistungsfähigkeit durch die Automation selbst. Sie können situativ eine Erwartung möglicher kritischer Systemzustände und Systemverhaltens erzeugen und dadurch das Automationsbewusstsein als auch die Akzeptanz der Automation erhöhen.

Da aber entsprechende Kontrollübernahmezeiten, je nach Fahrerablenkung, mehrere Sekunden dauern können (z.B. Damböck, 2012; Gold & Bengler, 2014, Petermann 2013), müsste auch die Automationsrückmeldung entsprechend mehrere Sekunden vor der eigentlich problematischen Situation geschehen. Zeitspannen in denen die Vorhersage zukünftiger Systemzustände bzw. möglicher Probleme der Automation in einer Situation durch die Automation sehr unzuverlässig werden kann.

Alarmsysteme, die mittels Rückmeldungen situativ eine Erwartung potentiell negativer Situationen erzeugen, stoßen hier schnell an bestimmte Grenzen; nämlich, wenn aufgrund der unzuverlässigen Vorhersage die relative Anzahl falscher Alarme ansteigt. Das häufige Erleben falscher Alarme erzeugt einen Lerneffekt, der dazu führt, dass man sich nicht mehr auf den Alarm verlässt und diesen sogar ignoriert (z.B. Bresnitz, 1984, Bliss & Dunn, 2000). Dieser Effekt wird Cry Wolf Effekt genannt. Der Cry Wolf Effekt geht so weit, dass selbst in Hochrisiko-Situationen niedrigere Unterstützungsgrade des Alarmsystems gewählt werden oder es gar ausgeschaltet wird, wenn diesem nicht mehr vertraut wird (Cafarelli & Hansmann, 1998).

Klassische Alarmsysteme haben einen weiteren Nachteil, dass durch sie selbst nicht direkt nachvollziehbar wird, warum ein Eingriff erforderlich wird, sie liefern keine Begründung warum eine kritische Automationsaktion auftreten kann. In Arbeiten von Bagheri & Jamieson (2004) konnten diese z.B. zeigen, dass Kontextwissen zu möglichen Gründen von Automationsfehlverhalten die Kontrollierbarkeit von Fehlern und auch die Akzeptanz fehlerhafter Automation verbessern konnte. Konventionelle Alarmsysteme sind daher wahrscheinlich nicht geeignet, zuverlässig und zeitlich stabil, zu einem adäquaten Automationsbewusstsein in problematischen Fahrsituationen beizutragen. Im Gegenteil könnten häufige Übernahmeaufforderungen durch unzuverlässige Vorhersagesysteme nicht nur die

Effektivität einer solchen Rückmeldung selbst deutlich negativ beeinflussen, sondern wahrscheinlich auch insgesamt den wahrgenommenen Nutzen einer Fahrzeugautomation schmälern.

Es stellt sich in diesem Zusammenhang die Frage, ob es wirkungsvolle Rückmeldesysteme geben kann, die, trotz der unzuverlässigen Vorhersage zukünftiger Automationszustände, den Aufbau eines adäquaten Automationsbewusstseins ermöglichen.

Die Forschung zum Bereich des Cry Wolf Effektes konnte zeigen, dass die Rückmeldung von Automationsunsicherheit bzw. von wahrscheinlichkeitsbasierten Warnungen, z.B. bei sogenannten Likelihoodalarmen, den Cry Wolf Effekt teils deutlich minimieren kann (Sorkin & Kantowitz, 1988, Wiczorek & Manzey, 2011, Bisantz, Marsiglio & Munch, 2005, Bustamante, 2008, Finger & Bisantz, 2002, McGuirl & Sarter, 2006).

Die Anwendung von Unsicherheitsrückmeldung im Kontext des hochautomatisierten Fahrens wurde jedoch bisher nur ansatzweise untersucht. Im Kontext des automatisierten Fahrens lag dem Autor diesbezüglich nur eine Arbeit vor (Seppelt, 2009). Weitere Arbeiten entstanden erst nach der Veröffentlichung der ersten Studie der vorliegenden Arbeit (Helldin et al. 2013).

Ebenso wurde in keiner dem Autor bekannten Arbeit eine mögliche zugrundeliegende Theorie der Wirkung einer Unsicherheitsrückmeldung untersucht. Als erklärende Theorie der Effektivität von Unsicherheitsrückmeldung bei häufigen falschen Alarmen wird in der vorliegenden Arbeit die Theorie der Evidenzevaluation, einer Theorie des kausalen Lernens (z.B. White, 2002), betrachtet. Nach dieser stellen falsche Alarme negative Evidenzen für eine Zusammenhangshypothese zwischen Alarmursache und Alarm dar. Wenn zu viele negative Evidenzen erlebt werden, sinkt die Stärke des vermuteten Zusammenhangs. Eine Unsicherheitsrückmeldung hingegen lässt den Ausgang offen. Das Eintreten bzw. Nicht-Eintreten einer Situation, kann nicht als positive bzw. negative Evidenz für das Vorliegen von Unsicherheit gezählt werden, sondern beide Zustände sind mögliche und eventuell gleichberechtigte Konsequenzen einer unsicheren Vorhersage. Ein weiterer potentieller Vorteil einer Unsicherheitsrückmeldung ist der, dass der Status „Automation ist unsicher“ direkt als eine Art Erklärung für ein der Fahrsituation nicht angemessenes Verhalten der Automation gesehen werden kann; nämlich, weil die Automation unsicher ist, obwohl die Gründe für die Automationsunsicherheit selber gar nicht kommuniziert werden. Das Vorhandensein einer Begründung für ein, aus dem Auge des Fahrers, mögliches fehlerhaftes Automationsverhaltens, könnte die Akzeptanz einer solchen nicht-perfekten Automation erhöhen.

## *1.1 Begriffsklärung Automationsfehler & falsche Alarme*

In dieser Arbeit wird immer wieder von Automationsfehlern gesprochen. Damit sind Situationen gemeint, in denen sich die Handlungen der Automation, ohne das Eingreifen eines Fahrers, zu einer Gefahr entwickeln können. Häufig sind dies strenggenommen, keine Fehler der Automation, da die Automation für diese Situationen eben nicht entwickelt wurde, sich also mitunter außerhalb der Systemgrenzen liegt. Ein echter Automationsfehler definiert sich nur durch ein falsches Verhalten innerhalb der Systemgrenzen. Da dies einem Fahrer aber nicht immer klar sein dürfte, warum eine Automation nicht angemessen reagiert, und eben Systemgrenzen häufig nicht klar sind, stellt sich das Verhalten dem Fahrer als Automationsfehler dar. Daher wird der Begriff Automationsfehler genutzt, auch wenn es sich tatsächlich um Automationsverhalten außerhalb von Systemgrenzen handelt. In dieser Arbeit werden grundsätzlich zwei Teilsysteme unterschieden:

- das automatische Teilsystem, welches die Fahraufgabe übernimmt und die oben genannten Automationsfehler zeigen kann
- das automatische Teilsystem welches vor solchen Automationsfehlern warnen soll. Ein solches System wird als Alarmsystem bzw. auch allgemeiner als Rückmeldesystem betrachtet. Fehler dieses Systems werden in dieser Arbeit nicht als Automationsfehler, sondern als falsche Alarme bezeichnet.

## *1.2 Ziele der Arbeit*

Ziel dieser Arbeit ist die Untersuchung einer Unsicherheitsrückmeldung, als möglichem Baustein zur Verbesserung des Automations-/Situationsbewusstseins und der Kontrollierbarkeit von Automationsfehlern, beim Teil- bis hochautomatisiertem Fahren.

- Es sollen Aussagen zur praktischen Relevanz eines Konzeptes zur Unsicherheitsrückmeldung für zukünftige Formen des automationsunterstützten Fahrens ermöglicht werden, in wie weit sie effektiv und effizient zur Erhöhung der Sicherheit des teil- bis hochautomatisierten Fahrens ist.
- Dazu wird insgesamt betrachtet, wie sich Unsicherheitsrückmeldung im Vergleich zu anderen Rückmeldungen auf die situative Steigerung der Erwartung von Automationsfehlern auswirkt, als auch wie sich eine Unsicherheitsrückmeldung auf die allgemeine Erwartung von Automationsfehlern in erster Linie durch Anpassung des Vertrauens in die Automation auswirkt.

- Es wird ein Schwerpunkt auf die Untersuchung gelegt, ob durch die Art der Alarmgestaltung als Unsicherheitsrückmeldung eine Anfälligkeit für einen potenziellen Cry Wolf Effekt minimiert werden kann. Dabei wird ein lerntheoretischer Erklärungsansatz für die Auswirkungen unterschiedlicher Alarmgestaltungen auf die Anfälligkeit einen Cry Wolf Effekt zu entwickeln genutzt.

## **2 Theoretischer Hintergrund**

### *2.1 Fahrzeugautomation und die Rolle des Fahrers*

Diverse Unternehmen, darunter die klassischen Automobilhersteller aber auch „Tech“ Unternehmen wie Waymo (Google), Uber und Weitere, testen hochautomatisiertes Fahren schon seit geraumer Zeit im Straßenverkehr. Abgesehen von einigen noch deutlich herausfordernden Situationen, wie sie z.B. im Innenstadtverkehr auftreten, steht hochautomatisiertes und autonomes Fahren in vielen Situationen vor keiner besonderen technischen Hürde mehr. Zu den noch existierenden Hürden sind vor allem die Rechtlichen zu zählen. Hier ist es insbesondere die Frage der Haftung, die es zu klären gilt, wenn ein durch eine Automation unterstütztes Fahrzeug einen Unfall verursacht, bzw. in einen Unfall verwickelt ist.

Um die Frage der Verantwortung bzw. der Haftbarkeit besser beantworten zu können, wurde im Auftrag der Bundesanstalt für Straßenwesen (BASt) von einer Expertenkommission definiert, welche Bedingungen erfüllt sein müssen, damit verschiedene Ausprägungen einer Fahrzeugautomation umgesetzt werden können (2012 (Gasser et al)). Im Endeffekt kann die Verantwortung für die Fahrzeugkontrolle immer nur entweder beim Fahrer oder bei der Fahrzeugautomation liegen. Es wurden insgesamt 5 Automationsstufen und entsprechende Bedingungen definiert, die sich durch eine unterschiedliche Involvierung des Fahrers in die Ausführung der Fahraufgabe ergaben. Diese 5 Stufen unterscheiden sich vor allem dadurch, wie und wann die Kontrolle über die Fahraufgabe eindeutig an entweder den Fahrer oder die Automation verteilt wird.

- Beim „manuellen Fahren“ fährt der Fahrer vollständig selbständig. Eine Automation übernimmt keinen wesentlichen Teil der Fahraufgabe. Die Verantwortung liegt dauerhaft beim Fahrer.
- Beim „assistierten Fahren“ übernimmt eine Automation entweder die Längsführung oder die Quersführung, der Fahrer entsprechend die andere Führungsdimension. Der Fahrer ist also permanent aktiv in die Ausführung der

Fahraufgabe eingebunden. Durch die permanente aktive Involvierung liegt die Verantwortung zu 100% beim Fahrer.

- Beim „teilautomatisierten Fahren“ übernimmt die Automation sowohl die Längs- als auch die Querführung. Der Fahrer muss aber immer passiv in die Fahraufgabe eingebunden sein, um jeder Zeit, ohne Vorwarnzeit, die Fahraufgabe übernehmen zu können. Auch hier liegt die Verantwortung zu 100% beim Fahrer.
- Beim „hochautomatisierten Fahren“ übernimmt die Automation wieder die Längs- und Querführung, der Fahrer darf sich mit Nebenaufgaben beschäftigen, muss aber in der Lage sein nach einer bestimmten Vorwarnzeit die Fahraufgabe wieder zu übernehmen. Nach dieser Vorwarnzeit liegt die Verantwortung wieder zu 100% beim Fahrer.
- Beim „autonomen Fahren“ übernimmt die Automation die Längs und Querführung in definierten Anwendungskontexten (z.B. Autobahnfahrt) und der Fahrer muss in keiner Weise mehr in die Ausführung der Fahraufgabe eingebunden sein. Das Fahrzeug muss selbst in der Lage sein, einen sicheren Zustand zu erreichen. Hier liegt die Verantwortung tatsächlich bei der Fahrzeugautomation. Verlässt das Fahrzeug den Anwendungskontext muss der Fahrer mit einer ausreichenden Zeitreserve die Fahraufgabe übernehmen.

Weitere Organisationen wie die Society of Automotive Engineers (SAE, 2014) und die US National Highway Traffic Safety Administration (NHTSA, 2013) haben ähnliche Einstufungen herausgebracht, die sich vor allem in den unteren Automationsgraden nicht unterscheiden – bei höherer Automation, insbesondere aber beim autonomen Fahren, über die BAST Definitionen hinausgehen und z.B. keine Limitierung mehr auf bestimmte Anwendungskontexte sehen.

Allen Stufen, mit Ausnahme des autonomen Fahrens, ist gemein, dass der Fahrer entweder permanent in die Ausführung der Fahraufgabe involviert sein muss, um jeder Zeit die Kontrolle sofort übernehmen zu können, oder in einer relativ kurzen Zeit wieder in einem ausreichenden Maß in die Fahraufgabe involviert werden muss. Erst im autonomen Fahren darf sich der Fahrer vollständig aus der Fahraufgabe zurückziehen. Hier muss die Funktionssicherheit durch die Hersteller in einem solchen Maß gewährleistet sein, dass der Fahrer dem Fahrzeug bzw. der Automation auch berechtigt vertrauen kann. Dass aktuell die Fahrzeugautomation bei allem Fortschritt noch nicht so weit ist, zeigten unlängst einige tödliche Unfälle in die teilvollautomatisierte Fahrzeuge verwickelt waren. (Uber, Tesla etc. siehe z.B. [https://en.wikipedia.org/wiki/List\\_of\\_self-driving\\_car\\_fatalities](https://en.wikipedia.org/wiki/List_of_self-driving_car_fatalities).)

## 2.2 Fahrzeugautomation und Situationsbewusstsein

Die Involvierung des Fahrers in die Fahraufgabe ist notwendig, um ein der Situation angemessenes Situationsbewusstsein aufzubauen und aufrechtzuerhalten. Ein Bewusstsein über die aktuelle Situation ist wiederum die Grundlage für ein der Situation angemessenes Verhalten. Das am häufigsten zitierte Modell stammt von Endsley (1995). Nach Endsley (1995) ist Situationsbewusstsein:

*“Situation awareness is the **perception** of the elements in the environment within a volume of time and space, the **comprehension** of their meaning and the **projection** of their status in the near future“ (Endsley, 1995)*

Es handelt sich demnach um ein 3 stufiges Modell mit folgenden Stufen:

1. Stufe - Wahrnehmung: Informations-extraktion aus der Umgebung
2. Stufe - Bedeutung: Informations-integration zu einem mentalen Bild
3. Stufe - Projektion: Vorausahmen zukünftiger Ereignisse

Endsley sieht dabei das Situationsbewusstsein selbst als Zustand – der durch den Prozess des „Situation Assessment“ erreicht wird. Bei einigen anderen Autoren wird Situationsbewusstsein aufgrund der Dynamik der Umwelt eher als Prozess gesehen (z.B. Adams, Tenney & Pew, 1995 oder Smith & Hancock, 1995).

Situationsbewusstsein ist dabei ebenfalls das Produkt eines mehrstufigen Prozesses, der ähnlich wie in Endsleys Modell, die Stufen Informationsextraktion, Informationsintegration und das Vorausahmen zukünftiger Verhaltensweisen beinhaltet. Es wird aber nach der Informationsintegration noch eine Zwischenstufe eingefügt – das gezielte Lenken der Aufmerksamkeit aufgrund der Erkenntnisse des Integrationsprozesses (Stufe 2) (Dominguez, 1994)

## 2.3 Automationawareness – Erwartung des Automationsverhaltens

Nach einer Definition von Wickens (1996) wird das Situationsbewusstsein explizit nicht nur auf die äußere Situation bezogen, er bringt den Begriff des dynamischen Systems ins Spiel.

*“The continuous **extraction of information** about a dynamic system or environment, the integration of this information with previously acquired knowledge to form a coherent **mental picture**, and the use of that picture in directing further perception of, anticipation of, and attention to **future events**“ (Wickens, 1996)*

Mit dynamischen Systemen sind dabei in erster Linie Automationssysteme bzw. Mensch-Maschine Systeme gemeint, die genauso im Gesamtbild eines Situationsbewusstseins berücksichtigt sein müssen, wie Gegebenheiten der Umwelt. Explizit wird dies z.B. von Sarter & Woods (1995) als „Mode Awareness“ oder Automation Awareness beschrieben:

*“Mode awareness is the ability of a supervisor to track and to anticipate the behavior of automated systems!” (Sarter & Woods, 1995)*

Das Situationsbewusstsein über die Automation und ihre Leistungsfähigkeit, also auch die Antizipation des korrekten Verhaltens einer Automation in der Zukunft – sind als wesentliche Voraussetzung zu sehen, ein ausreichend hohes Situationsbewusstsein einer Gesamtsituation beim Fahren mit einer Automation zu erlangen.

Insbesondere in der Interaktion mit einer Automation ist es schwierig, wie in unterschiedlichen Studien gezeigt wurde, ein adäquates Situationsbewusstsein aufzubauen, bzw. aufrechtzuerhalten, selbst wenn vom Fahrer gefordert wird, dass dieser permanent in die Ausführung der Fahraufgabe involviert ist.

## **2.4 Situationsbewusstsein & Automationsgrade**

Die in der Definition des teilautomatisierten Fahrens zur Aufrechterhaltung eines ausreichenden Situationsbewusstseins erforderliche permanente Involvierung in die Fahraufgabe ist eher eine theoretische Möglichkeit, tatsächlich auf die Übernahme der Fahraufgabe vorbereitet zu sein. In der Praxis zeigt sich, dass z.B. Menschen aus mehreren Gründen keine Veranlassung sehen, dieser Überwachung nachkommen zu müssen. Diesen Schluss lassen die Ergebnisse einer Reihe von Untersuchungen zu.

Rudin Brown & Parker (2004) untersuchten in einer Studie mit Realfahrzeugen die Auswirkungen einer automatisierten Längsführung (Adaptive Cruise Control - ACC) auf die Leistung in einer Zweitaufgabe und einer weiteren Aufgabe, in der die Fahrer Gefahren entdecken sollten. Es wurden unterschiedliche Ausprägungen der Automation miteinander verglichen (kein ACC und 2 ACC Varianten mit unterschiedlichem Abstand zum Vorderfahrzeug). Es zeigte sich, dass in den ACC Varianten zwar mehr Zweitaufgaben bearbeitet, dafür aber weniger Gefahren entdeckt wurden.

De Waard et al (1999) untersuchten in einer Simulatorstudie teilautomatisiertes Fahren in Kolonnen in denen Fahrzeuge mit vergleichsweise niedrigem Abstand



zueinander fahren (1 Sekunde Time Headway und 0,25 Sekunden). Es zeigte sich hierbei, dass von den Fahrern Notfallsituationen und Systemgrenzen, in denen sie zur Vermeidung eines Unfalls die Längsführung übernehmen mussten, nicht erwartet wurden. Dies äußerte sich darin, dass nur die Hälfte aller Versuchspersonen eine Kollision in einer solchen Situation vermeiden konnte.

Kastner & Baumann (2011) untersuchten in einer Simulatorstudie den Einfluss assistierten Fahrens mit einem Längsführungsassistenten – Adaptive Cruise Control (ACC) – im innerstädtischen Stop & Go Verkehr. Die Probanden fuhren darin, entweder mit oder ohne ACC Unterstützung, einem Fahrzeug hinterher. Es konnten in der Bedingung assistiertes Fahren Mitzieheffekte beobachtet werden, in denen Probanden eine rote Ampel nicht oder deutlich zu spät bemerkten und diese dadurch unberechtigt überquerten.

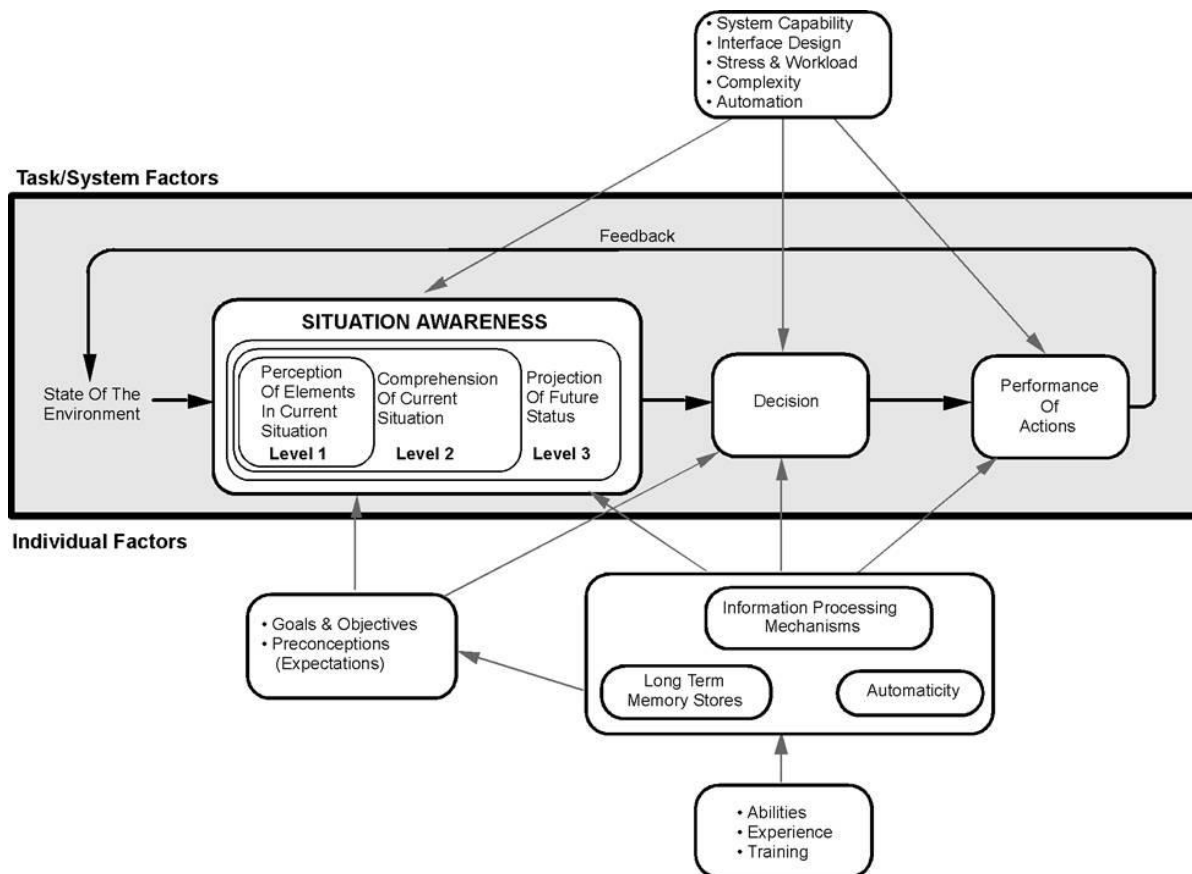
Neukum et al. (2008) untersuchten in einer Realverkehrsstudie Reaktionen von Fahrern auf das Beschleunigen eines mit einem Längsführungsassistenten ausgestatteten Fahrzeug auf ein stehendes Hindernis, nachdem das Assistenzsystem vorher das Zielobjekt verloren hatte. Die Probanden zeigten hier verspätete Bremsreaktionen. Vor allem subjektiv wurden die Ereignisse als deutlich kritisch eingeschätzt. Ein Hinweis darauf, dass nicht angepasstes Systemverhalten an Systemgrenzen nicht erwartet, bzw. vorhergesehen wird.

All diese Untersuchungen können in unterschiedlichen Bereichen zeigen, dass sich Fahrzeugautomation, selbst in Fällen, in denen der Fahrer noch recht deutlich und aktiv in die Ausführung der Fahraufgabe eingebunden ist, so negativ auf das Situationsbewusstsein auswirkt, dass Systemgrenzen nicht vorhergesehen werden und Fahrer teils ungenügend auf Systemausfälle reagieren.

## *2.5 Falsche Erwartung des Automationsverhaltens - Overtrust & Complacency*

Nach Endsleys Modell des Situationsbewusstseins (1995), wirken sich verschiedene Prozesse bzw. Zustände auf die Bildung des Situationsbewusstseins aus.

Wesentliche davon sind die bisherige Erfahrungen mit (Automations-) Systemen – bzw. die Erwartung wie sich diese in bestimmten Situationen verhalten werden (siehe Abbildung 1).



**Abbildung 1: Die Stufen des Situationsbewusstseins an der Entscheidungsfindung und die an der Entstehung des Situationsbewusstseins beteiligten Prozesse nach Endsley (1995)**

Die Erwartung darüber wie sich ein Automationssystem in einer bestimmten Situation verhalten wird, wird maßgeblich durch das sogenannte Vertrauen in die Automation beeinflusst.

Übermäßiges Vertrauen (Trust) in die Automation und die daraus resultierende Complacency Problematik sind wesentlichen Gründe, warum der Mensch der Forderung nach einer permanenten Überwachung nicht nachkommen kann.

Trust in Automation (z.B. Masaloni & Parasuraman, 1999; Lee & See, 2004) und dadurch verursachte sogenannte Complacency (z.B. Parasuraman, Molloy & Singh, 1993) beschreibt die Neigung des Menschen komplexen technischen Systemen zu Vertrauen, wenn sie aufgrund mangelnder Informationslage, oder einfach aufgrund ihrer schieren Komplexität nicht durchschaut werden können. Vertrauen bedeutet in diesem Zusammenhang die implizite Erwartung, dass sie bestimmte Situationen mit ausreichender Sicherheit bewältigen können, ohne dies logisch oder durch technisches Systemverständnis herleiten zu können.

Dieses Vertrauen scheint umso größer zu sein, je mehr Handlungen die Automation übernimmt bzw. je höher die Automatisierung ist. Diese impliziten Annahmen stehen dem expliziten Verständnis einer Automation gegenüber, dem Wissen eines Nutzers einer Automation über die Funktionsweise einer Automation & deren Systemgrenzen.

Explizites Wissen entspricht dabei dem sogenannten mentalen Modell (z.B. Johnson-Laird, 1983) welches ein Nutzer über eine Automation bzw. eine Technik hat.

Übervertrauen liegt dann vor, wenn trotz unzureichender Erfahrung angenommen wird, dass die Automation deutlich mehr Situationen bewältigen kann bzw. diese deutlich besser bewältigen kann, als sie tatsächlich in der Lage ist. Es handelt sich demnach um falsche Erwartungen aufgrund bestimmter Vorurteile (Preconceptions) die in recht kurzer Auseinandersetzung mit dem Automationssystem gebildet werden. Nach der weiter oben beschriebenen Mode- bzw. Automation Awareness, führt ein nicht angepasstes Systemvertrauen zu einer mangelnden Integration relevanter Hinweise auf ein zukünftiges Automationsverhalten (SA Stufe2), sofern solche Hinweise überhaupt existieren bzw. erkannt werden können. Folge dessen ist eine mangelnde Antizipation (SA Stufe 3) möglicher Systemgrenzen.

Complacency beschreibt das aus Übervertrauen resultierende Verhalten, der Vernachlässigung der eigentlich aufgetragenen Pflichten, nämlich der permanenten Überwachung (z.B. Parasuraman, Molloy & Singh, 1993).

### 2.5.1 Trust und Complacency bei Fahrzeugautomation

Im Kontext des Fahrens gibt es einige Untersuchungen zu Auswirkungen von Trust und durch Trust erzeugter Complacency. Niederee & Vollrath (2009) z.B., untersuchten in einer Fahrsimulatorstudie wie Probanden mit Systemausfällen bei unterschiedlichen Automatisierungsstufen in Form von Geschwindigkeitswarnung und Adaptive Cruise Control umgingen. Es wurde systematisch der Einfluss unterschiedlicher Zuverlässigkeit der Assistenzsysteme erhoben. Es konnte ein Zusammenhang beobachtet werden, in denen Versuchspersonen umso schlechtere Reaktion zeigten, je höher und je zuverlässiger die Automation im Laufe der erlebten Fahrt gewesen ist. Dies ist ein Hinweis darauf dass eine niedrige Einbindung des Fahrers in die Fahraufgabe ein schlechtes Situationsbewusstsein fördert und insbesondere bei anscheinend zuverlässigen Systemen ein Systemausfall nicht vorhergesehen wird. Die Untersuchung zeigte aber auch, dass das Erleben einer deutlich fehlerhaften Automation zu einer Verbesserung der Leistung der Kontrollübernahme führte – hier also eine Trustanpassung durch Erfahrung geschah.

Bahner (2008) untersuchte in zwei Studien im Kontext der Prozessautomatisierung den Einfluss von Fehlererfahrungen auf die Erwartung von Automationsfehlern (automation bias), also auch auf die Anpassung des Trust Niveaus hin zu einem der tatsächlichen Automationsleistungsfähigkeit angepassten Trust. Versuchspersonen interagierten hier mit einer Automation, die sie bei der Detektion, Diagnose und Behebung von Fehlern in der Prozesssteuerung unterstützen sollte.

In beiden Studien wurde der Einfluss der Erfahrung von Automationsfehlern unterschiedlicher Art (Fehldiagnosen und Ausfälle) in einem Trainingsdurchlauf auf die Complacencyneigung in einem späteren Durchgang erfasst. Bahner konnte beobachten, dass sich die unterschiedlichen Erfahrungen spezifisch auf die Interaktion mit der Automation und die Erwartung von Fehlern auswirkten. Allgemein konnte aber beobachtet werden, dass das Erleben von Automationsfehlern während des Trainings zu einer geringeren Complacency in Folgedurchläufen führte. Erfahrung von echten Automationsfehlern kann in der Realität jedoch nur schwer gesammelt werden, da die Basisrate tatsächlicher Fehler sehr niedrig ist. Tatsächliche Fehler treten bei einer Automationsfahrt zu selten auf, oder werden gar nicht wirklich wahrgenommen (z.B. Kraiss, 1994). Dies führt zu einer Verzerrung in Richtung des Übervertrauens und somit Complacency.

## *2.6 Kontrollübernahmezeiten*

Im Gegensatz zum teilautomatisierten Fahren wird in der Definition des hochautomatisierten Fahrens dem Fahrer zugestanden, sich weitgehend aus der Fahraufgabe zurückzuziehen. Der Fahrer darf sich aber nur unter der Voraussetzung zurückziehen, dass er vor einem möglichen Eingriff früh genug wieder in die Ausführung der Fahraufgabe zurückgeholt wird, um ein ausreichendes Situationsbewusstsein bis zum Zeitpunkt der tatsächlichen Kontrollübernahme erwerben zu können. Dies setzt voraus, dass das Automationssystem selbst weiß, dass es eine Kontrolle nicht sicher aufrechterhalten kann und den Fahrer durch Rückmeldung dazu bringt, sich wieder der Ausführung der Fahraufgabe anzunehmen. Dabei ist ein wesentlicher Aspekt, dass die Automation selbst aktiv den Fahrer auf die Kontrollübernahme hinweist, und zwar vor jeder möglichen Systemgrenze. Der Fahrer muss hier nicht mehr selbst aus Situationsmerkmalen schließen und antizipieren, wann eine Automation an einer Grenze angekommen ist. Die Automation fördert durch Systemfeedback die Bildung einer Automation-Awareness bezüglich der anstehenden notwendigen Kontrollrücknahme und dadurch die Hinwendung der Aufmerksamkeit zur Umwelt zwecks Bildung eines angemessenen Situationsbewusstseins.

Bei dieser Art des Zurückholens des Fahrers in die Fahraufgabe, ist es wesentlich zu wissen wie lange es unter verschiedenen Einflussfaktoren dauert, ein genügend großes Situationsbewusstsein aufzubauen, um die Kontrolle in ausreichender Qualität übernehmen zu können.

Es existieren einige Untersuchungen, die dieser Frage nachgehen.

Damböck, Farid, Töner & Bengler (2012) untersuchten in einer Fahrsimulatorstudie mit 32 Versuchspersonen, die Zeit, die es benötigt, die Kontrolle über ein Fahrzeug aus dem hochautomatisierten Fahren wieder zu erlangen, nachdem dieses in verschiedenen Szenarien an eine Systemgrenze kam. Diese unterschiedlichen Szenarien sollten die 3 Ebenen des Handlungsmodells nach Rasmussen (1986:) fertigkeitsbasiert, regelbasiert und wissensbasiert widerspiegeln, und stellen somit eine Variation des „Involvierungsgrades“ dar den es benötigt, um in bestimmten Situationen die Kontrolle adäquat zurückzuerlangen zu können. Die fertigkeitsbasierte Ebene beschreibt dabei hochtrainierte, nicht bewusst ablaufende Handlungen, die demnach auch kein ausgeprägtes Situationsbewusstsein, zumindest nicht in der Stufe 3 (Antizipation), erfordern. Regelbasierte Handlungen sind auch zu einem hohen Grad trainiert, laufen aber größtenteils nach bewusstem befolgen bestimmter Regeln ab (z.B. Verkehrsregeln). Die wissensbasierte Ebene erfordert das Erschließen bzw. Antizipieren bestimmter Zustände, um entsprechend adäquate Handlungen abzuleiten.

In der Studie entsprach Szenario 1 (fertigkeitsbasiert) dem Verlust von Fahrstreifenmarkierungen und erforderte ein Eingreifen auf der Stabilisierungsebene (Querführung). Situation zwei (regelbasiert) erforderte einen Fahrstreifenwechsel aufgrund einer Baustellensituation und Situation drei (wissensbasiert) verlangte eine Entscheidung des Fahrers, da sich an einer Gabelung zwei Richtungsmöglichkeiten auftaten. Während der Fahrt bearbeiteten die Fahrer eine Nebenaufgabe. In der Situation selbst fiel dann die Automation aus, angekündigt durch einen auditiven Hinweisreiz.

Es zeigte sich eine Abhängigkeit der Kontrollübernahmezeiten von der Komplexität der Fahrsituation. Auf der fertigkeitsbasierten Ebene lag die Übernahmedauer bei etwa 4 Sekunden. In den beiden anderen Szenarien zeigte sich eine Übernahmezeit von etwa 6 Sekunden: aber auch nach diesen 6 Sekunden machten die Fahrer deutlich mehr Fehler als in einer Vergleichsbedingung ohne Fahrzeugautomation. Die Autoren schlossen das in solch komplexeren Situationen eine Übernahmezeit von etwa 8 Sekunden nötig ist, um eine ähnlich gute Leistung wie beim manuellen Fahren zu erreichen.

Petermann-Stock, Hackenberg, Muhr & Mergl (2013) untersuchten in einem Fahrsimulator die Auswirkungen unterschiedlicher Nebenaufgaben auf die Übernahmezeit bei einer hochautomatisierten Fahrt mit einem Stauassistenten. Die Nebenaufgabe wurde nach Höhe und Art der Beanspruchung variiert. Im niedrigen Beanspruchungslevel mussten die Probanden während der Fahrt Fragen beantworten, die ihnen der Versuchsleiter stellte. Die Beanspruchung ist hier kognitiv und auditiv. Im mittleren Beanspruchungslevel mussten die Probanden wieder

Fragen beantworten, die der Versuchsleiter stellte. Diesmal wurden die Antwortmöglichkeiten jedoch auf einem Bildschirm dargeboten und die korrekte Antwort musste vorgelesen werden. Die Art der Beanspruchung ist hier also auditiv, visuell und kognitiv. In der Stufe hohe Beanspruchung wurden die Fragen wieder durch den Versuchsleiter gestellt, der Proband musste die richtige Antwort auf einem Bildschirm identifizieren und anschließend über eine Tastatur eintippen. Die Art der Beanspruchung hier ist auditiv, visuell, motorisch und kognitiv. Die genannten Beanspruchungen sind der Beanspruchung ähnlich die ein Fahrer natürlicherweise während Nebentätigkeiten beim hochautomatisierten Fahren haben könnte. Die Versuchspersonen bearbeiteten diese Nebenaufgaben während einer Fahrt über eine Autobahn. In einem Stauende, das zudem in einer Kurve lag, mussten die Fahrer die Kontrolle über das Fahrzeug hinter einem stark bremsenden Fahrzeug übernehmen. Die Automation übernahm dabei die Quer- und Längsführung bis 35 km/h. Danach mussten die Fahrer die Kontrolle innerhalb einer Sekunde übernehmen, angekündigt durch einen Ton und eine Aufforderung im Display. Taten sie das nicht, so ereignete sich nicht sofort ein Unfall, wenn sie jedoch mehrere Sekunden lang nicht eingriffen, konnte eine Kollision nicht vermieden werden. Es zeigte sich ein Effekt der Nebenaufgabe bzw. des Beanspruchungslevels auf die Übernahmezeit. Geordnet nach aufsteigendem Beanspruchungslevel ergaben sich mittlere Reaktionszeiten von 2,03 Sek, 2,90 Sek und 3,35 Sek. Es wurde eine maximale Übernahmezeit von 8,8 Sekunden beobachtet, ebenso berichteten die Probanden, dass sie sich eine deutlich frühere Übernahmeaufforderung gewünscht hätten. Die Autoren schließen, dass für eine Automation mindestens 5 Sekunden, maximal jedoch 10 Sekunden Übernahmezeit zur Verfügung stehen sollte.

Vogelpohl, Vollrath & Kühn (2016) untersuchten in einer Fahrsimulatorstudie den Einfluss von Müdigkeit auf die Zeit, die es benötigt, die Kontrolle über ein hochautomatisiertes Fahrzeug nach einer Übernahmeaufforderung zurück zu erlangen. Dabei fuhren Versuchspersonen so lange in einem Fahrsimulator bis diese anhand bestimmter Merkmale als müde eingeschätzt wurden. Darauf wurde eine Übernahmeaufforderung bestehend aus einem akustischen Hinweisreiz und einem Symbol im Display, das Hände am Steuer zeigte, ausgelöst. Es wurden eine Reihe von Reaktionsparametern erfasst (z.B. Zeit den Blickfokus eindeutig zur Straßenmitte zu lenken) und mit denen von Fahrern manueller Fahrzeuge verglichen, die zum gleichen Zeitpunkt eine Warnung bekamen. Es zeigte sich, dass hochautomatisiertes Fahren deutlich schneller zu Müdigkeit führt als das manuelle Fahren. So waren Fahrer hochautomatisierter Fahrzeuge etwa 15-20 Minuten früher müde als die Fahrer manueller Fahrzeuge. Es zeigte sich aber auch, dass eine hohe Müdigkeit beim hochautomatisierten Fahren zu einer höheren Übernahmedauer von etwa 3

Sekunden mehr im Vergleich zum manuellen Fahren führte. Insgesamt konnten Übernahmezeiten von 5-7 Sekunden beobachtet werden (Zeit bis die Automation abgeschaltet wurde). Bei anderen, für den Aufbau des Situationsbewusstseins wichtigen Parametern, wie etwa der Blick in den Spiegel etc. wurden sogar Zeiten von 12-15 Sekunden beobachtet.

Es existieren darüber hinaus weitere Studien, die sich mit Übernahmezeiten bei hochautomatisiertem Fahren beschäftigten. Eine sehr gute Übersicht über den bisherigen Literaturstand bietet Vogelpohl, Vollrath et al (2016). Hier wurden auch alle bisher identifizierten Übernahmezeiten in einer Tabelle zusammengefasst. Es zeigte sich, immer in Abhängigkeit der Verkehrssituation, der Ablenkung des Fahrers bzw. des Fahrerzustandes, der Bedingungen des Fahrzeuges aber auch der Art der Systemrückmeldung, eine zeitliche Variabilität von 1.16 Sekunden im Bestfall (Wulf et al. 2013) bis hin zu 35 Sekunden (Merat et al. 2014) oder gar mehreren Minuten (Brandenburg & Skottke, 2014) im schlechtesten Fall.

## *2.7 Die Problematik der „rechtzeitigen“ Rückmeldung*

### 2.7.1 Das Vorhersagedilemma – Signalentdeckungstheorie

Vogelpohl et al. (2016) schließen aufgrund des Literaturreviews und ihren eigenen Untersuchungen auf eine Zeit von 8 Sekunden, die es mindestens benötigt, um 90 % der Fahrer adäquat reagieren zu lassen. Dies sind 8 Sekunden, in der das Fahrzeug die Kontrolle vom Beginn der Übernahmeaufforderung bis zur tatsächlichen Übernahme zumindest in großen Teilen selbst ausführen muss.

8 Sekunden sind im Kontext hochdynamischer und komplexer Fahrsituationen eine verhältnismäßig lange Zeit, in der viel geschehen kann. So legt ein Fahrzeug bei beispielsweise 100 Km/h etwa 222 Meter zurück. Auch weitere Verkehrsteilnehmer, Infrastruktur wie Ampeln etc. können sich in dieser Zeit maßgeblich verändern. Aufgrund dieser hohen Dynamik kann es sein, dass der Grund der Alarmauslösung bzw. der Kontrollrückgabe nach diesen 8 Sekunden gar nicht mehr besteht, bzw. sich als deutlich weniger problematisch erweist, als durch eine Automation bzw. ein Alarmsystem antizipiert wurde.

Zur Erläuterung der Problematik kann die Signalentdeckungstheorie (bzw. Signal Detection Theorie - SDT) herangezogen werden (Green & Swets, 1966).

Die SDT beschreibt die Wahrscheinlichkeit, mit der ein Zielreiz in einer verrauschten Umgebung erkannt werden kann. Sie war ursprünglich auf die Fähigkeit eines Menschen bezogen, einen solchen Zielreiz entdecken zu können. Jedoch kann die SDT gut als Analogie zur Vorhersageproblematik einer Automation herangezogen

werden, da die maschinelle Wahrnehmung ebenfalls auf Wahrscheinlichkeitsberechnungen fußt (z.B. Darms, 2007). Als Rauschen wird in der SDT jeder Einflussfaktor, sowohl aus der Umgebung als auch innerhalb eines Automationssystems bzw. Detektionssystems (z.B. Rauschen im Sensor), gesehen der den eigentlich zu identifizierenden Reiz maskiert.

Das Antwortverhalten eines Systems kann in die Komponenten Sensitivität als auch Antworttendenz zerlegt werden. Die Sensitivität ist definiert als die Fähigkeit, zwischen den Zuständen Signal vorhanden und Signal ist nicht vorhanden unterscheiden zu können. Die Antworttendenz ist der Grad der Neigung, das Vorhandensein eines Signals zurückzumelden, wenn kein eindeutiger Zustand, Signal vorhanden oder Signal nicht vorhanden, vorliegt, also eine Unsicherheit besteht.

In unserer Umgebung ist fast jedes Signal verrauscht, da es kaum Reize gibt, die vollständig isoliert von anderen Einflussfaktoren auftreten. Insbesondere bei langfristigen Vorhersagen gibt es diverse Einflussfaktoren, die das Vorhandensein eines Signals (bzw. Ereignisses) in der Zukunft, maskieren können, ergo als Rauschen zu bezeichnen sind.

Bei jeder einzelnen Detektionsaufgabe gibt es laut der SDT folgende Ergebniszustände. Entweder ist ein Signal, bzw. ein berechtigter Grund für eine Kontrollrückgabe vorhanden und wird entdeckt – was als Treffer zu bezeichnen ist, oder es wird nicht entdeckt, obwohl es vorhanden ist. Dies wird als falsch negativ bezeichnet (siehe Tabelle 1). Solch falsch-negativen Zustände sind zu vermeiden, wenn es gelingen soll jeden tatsächlichen Automationsfehler kontrollierbar zu halten. Die anderen Fälle beziehen sich auf den Zustand des nicht Vorhandenseins eines Signals bzw. eines Grundes für eine berechtigte Kontrollrückgabe. Hier gibt es entweder eine korrekte Ablehnung oder ein Signal wird als vorhanden postuliert, obwohl es nicht vorhanden ist. Dies wird als falsch positiv bezeichnet, bzw. als falscher Alarm (siehe Tabelle 1). Die Implikationen, die solche falschen Alarme haben, werden weiter unten erläutert.

**Tabelle 1: Ergebnismöglichkeiten der Signalentdeckung nach der Signalentdeckungstheorie (Green & Swets, 1966)**

	Signal vorhanden	Signal nicht vorhanden
entdeckt	Treffer	Falsch Positiv (falscher Alarm)
Nicht entdeckt	Falsch Negativ (Miss)	Korrekte Ablehnung

Wie oben bereits beschrieben sollte es möglichst das Ziel sein, falsch Negative zu vermeiden. Dies geht aber in den meisten Fällen zu Lasten der falsch Positiven. Dies lässt sich Anhand der Theorie des Hypothesentestens verdeutlichen.



Häufig ist das Urteilen darüber, ob ein Signal vorhanden ist oder nicht, eine Frage der Wahrscheinlichkeit und die Entscheidung für eine bestimmte Antwort, basierend auf mehr oder weniger ausgeprägter Unsicherheit. Bei diesen Entscheidungen unter Unsicherheit werden Hypothesen formuliert, die gegeneinander getestet werden. Dabei geht man von der Nullhypothese aus „kein Signal ist vorhanden“. Ein falsch Positiver wäre somit ein Fehler erster Art (Alpha – Fehler) in dem die Nullhypothese abgelehnt wird. Das Vorhandensein eines Reizes wäre in diesem Kontext die Forschungshypothese und ein falsch Negativer wäre ein Fehler zweiter Art bzw. der Beta Fehler, die Nullhypothese beizubehalten obwohl die Forschungshypothese gilt (z.B. Macmillan, 2002).

Minimiert man den Alphafehler maximiert man zugleich den Betafehler – das heißt, will man möglichst alle Fälle, in denen ein Grund für eine Kontrollrückgabe tatsächlich vorhanden ist, abdecken, also falsche Negative vermeiden, so wird man recht zwingend die Anzahl der falsch Positiven erhöhen. In der Terminologie der SDT ist die Antworttendenz des Automationssystems liberal – also geneigt im Zweifelsfall für das Vorhandensein eines Reizes zu stimmen.

Bei sehr langen Vorhersagehorizonten ist entsprechend mit einer hohen Anzahl falscher Alarme bzw. Kontrollrückgaben zu rechnen (z.B. Parasuramann, Hancock & Olofinboba (1997).

### 2.7.2 Kontrollrückgabe oder dynamische Übernahmebereitschaft?

Ob es sich bei einer Kontrollrückgabe um einen falschen Alarm gehandelt hat, kann nur dann beurteilt werden, wenn der Fahrer nicht gezwungen wird die Kontrolle zu übernehmen, weil sonst z.B. das System ausgeschaltet würde. Der Fahrer kann durch die Übernahme der Kontrolle nicht überprüfen, ob die Automation in der Situation nicht doch richtig reagiert hätte.

Solche pauschalen Kontrollrückgaben können theoretisch negative Auswirkungen auf das Nützlichkeitsempfinden einer Automation haben und dazu führen, dass diese gar nicht benutzt wird, insbesondere dann, wenn in häufigen, für den Fahrer nicht nachvollziehbaren Gründen, die Kontrolle zurück an den Fahrer übergeben wird.

Dies wird in der Literatur „Disuse“ genannt (Parasuraman & Riley, 1997).

Eine andere Strategie wäre, die Kontrolle bei ersten Indizien auf eine Systemgrenze nicht gleich an den Fahrer zurückzugeben, ihn aber auf eine mögliche Übernahme vorzubereiten. Somit könnte die Zeit des unterstützten Fahrens möglichst lang und ununterbrochen gehalten werden. Der Fahrer könnte durch die Rückmeldung einer anstehenden Systemgrenze zurück in den „Loop“ geholt und übernahmebereit gemacht werden für eine mögliche Kontrollübernahme.

Hoc, Young & Blosseville (2009) bezeichnen in diesem Kontext die Automation als Partner (des Fahrers) in einem kooperativen System. Die Leistungsfähigkeit des

Systems kann durch das situative dynamische Anpassen der eigenen Leistung an die Leistungsfähigkeit des Automationspartners deutlich erhöht werden und das System insgesamt deutlich sicherer. Eine solch situativ bedingte dynamische Kontrollverteilung kann ein wesentlicher Bestandteil einer kooperativen Fahrzeugführung sein (z.B. Flemisch et al. 2012). Grundstein dieser dynamischen Kontrollverteilung ist die Kommunikation der aktuellen Leistungsfähigkeit der Automation an den (menschlichen) Partner.

Durch ein effizient gestaltetes Rückmeldesystem könnte situativ die Erwartung einer Systemgrenze bzw. damit verbundenem Automationsfehlverhalten erhöht werden, ohne gleich direkt die Kontrolle von der Automation an den Fahrer abzugeben.

Der Fahrer könnte somit ein dynamisches, der jeweiligen Situation angepasstes, Wechselspiel zwischen phasenweisem hochautomatisiertem Fahren ohne Überwachung der Automation und phasenweisen teilautomatisiertem Fahren mit Überwachung der Automation eingehen. In den Phasen des teilautomatisierten Fahrens ist der Fahrer alarmiert und übernahmebereit, in den Phasen des hochautomatisierten Fahrens kann der Fahrer sich weiteren Aufgaben widmen.

Nach dieser Strategie besteht jedoch die Gefahr, dass das die mögliche Folge eines „falschen Alarms“, also einer ausbleibenden Konsequenz nach einem Alarm direkt erlebt werden kann und somit das häufige Erleben falscher Alarme zu einem Absinken des Vertrauens in das Alarmsystem führt.

## *2.8 Alarmzuverlässigkeit und Vertrauen in das Alarmsystem*

Wickens & Dixon (2007) führten ein Literaturreview der Studien durch, welche sich mit den Auswirkungen unzuverlässiger Diagnosesysteme beschäftigte.

Diagnosesysteme sind hierbei alle Systeme, welche ein weiteres System überwachen (z.B. ein Subsystem in einem Flugzeug, um den Zustand der Treibwerke zu überwachen) und mit einem Alarm auf eine unzulässige Abweichung des zu überwachenden Systems von einem Sollzustand anzeigen sollen. Ziel der Studie war es, eine Aussage darüber treffen zu können, ab welchem Grad der Unzuverlässigkeit bzw. ab welchen Grad der falschen Alarme, eine Automation überhaupt noch einen Vorteil gegenüber dem Nicht-Vorhandensein eines solchen Diagnosesystems darstellen. Insgesamt flossen die Ergebnisse aus 22 Studien ein. Die Autoren kamen zum Schluss, dass es ab einem niedrigeren Niveau der Zuverlässigkeit als 0,70 die Kosten eines Diagnosesystems (z.B. durch die entstandene Ablenkung einem falschen Alarm nachzugehen) höher sind als der Nutzen. D.h., war die Automation in weniger als 70% der Fälle zuverlässig, wäre es besser gewesen, wenn der Nutzer das System gar nicht genutzt hätte.

### 2.8.1 Dimensionen des Vertrauens in ein Alarmsystem: Compliance & Reliance

Bezüglich möglicher Auswirkungen von Alarmsystemen mit niedriger Zuverlässigkeit postuliert Meyer (2001) zwei unterschiedliche Dimensionen des Vertrauens in ein Alarmsystem, die Compliance und die Reliance.

Compliance ist die Überzeugung, dass es tatsächlich eine Ursache für die Warnung gibt, wenn ein Alarm geäußert wird. D.h. bei einer hohen Compliance werden keine bis wenige false Positives erwartet. Reliance in das Alarmsystem hingegen ist der Grad der Überzeugung, dass das Alarmsystem anschlagen wird, wenn ein echter Grund für einen Alarm vorliegt. Bei einer hohen Reliance werden keine bis wenige false Negatives erwartet.

Dixon, Wickens & McCarley (2006) überprüften in einer Studie die These, dass Compliance und Reliance in ein Alarmsystem unabhängige Dimensionen sind. Dafür ließen sie Versuchspersonen in einem Luftfahrtszenario eine Trackingaufgabe gepaart mit einer Monitoringaufgabe durchführen. In der Monitoringaufgabe wurden die Versuchspersonen durch ein automatisches Diagnosesystem unterstützt. Die Autoren variierten die Anzahl der erlebten Fehler des automatischen Diagnosesystems in unterschiedlichen Bedingungen. In einer Bedingung produzierte es einen Anteil falscher Alarme und in einer anderen Bedingung Misses, also false Negatives. Die Autoren konnten beobachten, dass die Neigung des Systems falsche Alarme zu produzieren die Gesamtleistung der Versuchspersonen deutlich stärker negativ beeinflusste, als eine Automation, die dazu neigte, Misses zu produzieren. Das Erleben falscher Alarme wirkte sich sowohl negativ auf die Compliance als auch die Reliance aus. Die negative Auswirkung falscher Alarme auf die Reliance zeigte sich darin, dass die Leistung in der Trackingaufgabe deutlich abnahm, verglichen mit einer Vergleichsbedingung einer fehlerfreien Automation. Die Versuchspersonen schienen mit Fehlern zu rechnen, auch wenn gar kein Alarm geäußert wurde. Die Auswirkung auf die Compliance zeigte sich darin, dass insgesamt die Reaktionszeit auf die Alarme deutlich abnahm gegenüber der Bedingung mit der fehlerfreien Automation. Teilweise reagierten Versuchspersonen gar nicht mehr auf den Alarm. Misses wirkten sich hingegen nur auf die Reliance des Systems aus, was sich wieder in einer deutlichen Verschlechterung der Trackingaufgabe, verglichen mit der perfekten Automation zeigte. Das Erleben falscher Alarme führt also zu einer niedrigen Compliance in das Alarmsystem, was sich in niedrigeren Reaktionszeiten auf den Alarm und teilweise sogar ausbleibenden Reaktionen zeigt. Die Autoren schlossen darüber hinaus, dass das Erleben von falschen Alarmen tatsächlich negativere Auswirkungen auf das Operator Verhalten hat als das Erleben

von Misses, da auch in Bedingung „falscher Alarm“ dem Alarmsystem insgesamt weniger vertraut wurde.

### 2.8.2 Der Cry Wolf Effekt

In der Studie von Dixon, Wickens & McCarley konnte bereits beobachtet werden, dass sich falsche Alarme nicht nur in der Einstellungsebene, dem Vertrauen in ein Alarmsystem, niederschlagen, sondern auch deutlich auf der Verhaltensebene. Diese Neigung deutlich schlechtere Reaktionen auf einen Alarm als Folge des Erlebens falscher Alarme zu zeigen, wird in der Literatur false alarm bzw. häufig auch Cry Wolf Effekt genannt. (z.B. Bresnitz, 1984).

Bresnitz benannte den beobachteten Effekt nach einer Fabel von Aesop – in der ein Hirtenjunge mehrmals die Bewohner eines Dorfes zur Hilfe rief, da ein Wolf angeblich drohte seine Schafsherde anzugreifen. Diese kamen zur Hilfe, konnten aber niemals einen Wolf antreffen. Als dann tatsächlich ein Wolf die Herde anfiel, dachten die Dorfbewohner es würde sich wieder um einen falschen Alarm handeln und kamen diesmal nicht zur Hilfe.

Die Macht, mit der ein falscher Alarm zukünftige Erwartungen und Überzeugungen beeinflussen kann, findet seinen Ausdruck in der Alltagssprache, so z.B. auch im Sprichwort „Wer einmal lügt dem glaubt man nicht, und wenn er auch die Wahrheit spricht“ (Duden, 2002)

Auch andere Autoren konnten langsamere Reaktionszeiten auf Alarme eines Alarmsystems mit Neigung zu falschen Alarmen beobachten.

Getty et al. (1995) untersuchten in einem Laborexperiment die Auswirkungen unterschiedlicher Anteile falscher Alarme auf die Reaktionszeit. Dafür variierten die Autoren den Anteil falscher Alarme eines automatisierten Systems zur Detektion kritischer Zustände in einer Anlagen-Überwachungsaufgabe. Es gab die Stufen niedriger Anteil falscher Alarme (bzw. eine positive predictive value (PPV) – eine hohe Vorhersagegüte des Alarmsystems), mittlerer Anteil falscher Alarme (medium PPV) und hoher Anteil falscher Alarme (low PPV). Die Versuchspersonen mussten zeitgleich eine Trackingaufgabe durchführen. Mit einem Belohnungs– Bestrafungsparadigma wurden die Versuchspersonen motiviert, die Trackingaufgabe so akkurat wie möglich durchzuführen und zeitgleich gut auf Warnungen des Alarmsystems zu reagieren. Es zeigte sich, dass die Reaktionszeit auf Warnungen umso langsamer wurde, desto niedriger das PPV bzw. je höher die Anteile falscher Alarme waren. Getty et al. konnten ebenso wie Dixon, Wickens & McCarley beobachten, dass in einigen Fällen gar nicht mehr auf die Alarme reagiert wurde.

Bliss, Gilson & Deaton, 1995 – konnten in einer Untersuchung, die zum Ziel hatte die Existenz des Cry Wolf Effektes zu untermauern, beobachten, wie Versuchspersonen versuchten ihr Verhalten an die unterschiedlich erlebte Zuverlässigkeit eines Alarmsystems anzupassen – und sozusagen ihre Leistung in Abhängigkeit der Zuverlässigkeit des Alarmsystems zu optimieren. Auch hier wurde wieder ein Doppelaufgabenparadigma genutzt. Versuchspersonen mussten eine psychomotorisch fordernde Hauptaufgabe bearbeiten, während sie von einem Alarmsystem Alarme mit unterschiedlicher Dringlichkeit, niedrige, mittlere und hohe Dringlichkeit, präsentiert bekamen. Die Reliabilität des Alarmsystems wurde zwischen Gruppen variiert, in den Stufen 25% korrekte Alarme, 50% korrekte Alarme und 75% korrekte Alarme. Es konnte beobachtet werden, dass der Großteil der Versuchspersonen, etwa 90%, nicht immer auf Alarme reagierten, sondern ihre Reaktionsrate von der erlebten Häufigkeit falscher Alarme abhängig machten. Dieses Verhalten nannten die Autoren „Probability Matching“. Die Autoren schließen, dass Versuchspersonen ihr Verhalten an die subjektiv wahrgenommene Wahrscheinlichkeit eines korrekten Alarms anpassen. Etwa 10% aller Versuchspersonen wendeten eine Alles oder nichts Strategie an und reagierten entweder immer auf Alarme oder nie.

### 2.8.3 Untersuchungen zur Wirkung falscher Alarme im Kraftfahrzeug

Im Bereich des assistierten & automatisierten Fahrens sind bisher kaum Untersuchungen bekannt die sich mit den Auswirkungen falscher Alarme auf das Vertrauen in das Alarmsystem bzw. einen Cry Wolf Effect beschäftigen. Die zum Thema Cry Wolf Effekt existierenden Studien haben Kollisionswarnsysteme oder Notbremsassistenten zum Gegenstand.

Parasuramann, Hancock & Olofinboba (1997) beschäftigten sich auf theoretischer Ebene mit den Wahrscheinlichkeiten niedrigerer Zuverlässigkeit eines Frontal-Kollisionswarnsystems im Sinne häufiger falscher Alarme auf die Fahrerreaktionen. Sie betrachten hierfür wie in unterschiedlichen Fahrsituationen, mit unterschiedlichen Detektionsalgorithmen die Wahrscheinlichkeiten von Misses und False Alarms im Sinne der Signalentdeckungstheorie verringert werden können. Sie kommen zu dem Schluss, dass ein zufriedenstellendes Ergebnis wohl nicht erreicht werden kann. Bezüglich der Auswirkungen der falschen Alarme verweisen sie z.B. auf Erfahrungen aus dem Bereich der Luftfahrt mit sogenannten TCAS (Traffic Alert and Collision Avoidance System) Systemen, das bei Piloten häufig zu Missachtung und sogar Abschalten des Systems führt (z.B. Wiener, 1988). Der negative Effekt falscher Alarme führt hier sogar dazu, dass Systeme, die eigentlich nicht abschaltbar sind, mit

einer bestimmten Kreativität umgangen bzw. außer Kraft gesetzt werden. Satchell (1993) nutzte hierfür den Begriff des „Creative Disablements“.

Bliss & Acton (2003) untersuchten in zwei Simulatorstudien, mit jeweils 70 Versuchspersonen, wie sich unterschiedliche Grade der Alarmzuverlässigkeit eines Kollisionswarnsystems im Fahrzeug auf die Alarmreaktion des Fahrers auswirkt. Bei dem Kollisionswarnsystem handelte es sich nicht um ein Frontalkollisionswarnsystem, sondern um ein System, das vor Fahrzeugen warnte, die entweder von der Seite oder von hinten kamen und das Fahrzeug der Versuchspersonen nicht beachteten. In der ersten Studie wurde die Kollisionswarnung von der Mittelkonsole des Fahrzeuges aus geäußert. Das Alarmsystem hatte dabei drei verschiedene Zuverlässigkeitsstufen, 50%, 75% und 100% zuverlässig.

Im zweiten Experiment bekamen die gleichen Versuchspersonen Alarme mit der gleichen Zuverlässigkeit wie im ersten Experiment, diesmal wurden die Alarme jedoch nicht immer über die Mittelkonsole geäußert, sondern kamen immer aus der Richtung, aus der sich Fahrzeuge annäherten.

Unabhängig davon wie der Alarm geäußert wurde, also von der Mittelkonsole oder richtungsspezifisch, konnten Bliss & Action die gleichen Effekte beobachten, die Fahrerreaktionen waren signifikant besser je zuverlässiger die Alarme waren. Die Autoren kommen zu dem Schluss, dass Alarmsysteme möglichst eine hohe Zuverlässigkeit haben sollten.

## *2.9 Lerntheoretische Grundlagen des Cry Wolf Effektes*

Wenn die relative Häufigkeit von falschen Alarmen nicht verringert werden kann, besteht die einzige Möglichkeit einem Cry Wolf Effekt entgegenzuwirken darin, den Lernprozess des Fahrers bzw. des Alarmempfängers zu beeinflussen.

Um die Entstehung eines Cry Wolf Effektes systematisch beeinflussen zu können, mit dem Ziel dem Effekt entgegenzuwirken, ist es von Vorteil eine für die Erklärung des Effektes geeignete Lerntheorie zuordnen zu können. Erst wenn verstanden wird, welche Information durch welche Art des Lernens, wie berücksichtigt bzw. gewichtet und verarbeitet wird, kann darauf aufsetzend, untersucht werden, ob die Gestaltung der Alarmrückmeldung einen Effekt auf die Art des Lernens haben kann.

Trotz mannigfaltiger Literatur zum Cry Wolf Effekt bzw. des False Alarm Effektes, finden sich kaum Quellen welche mögliche zugrundeliegende Lernmechanismen thematisieren.

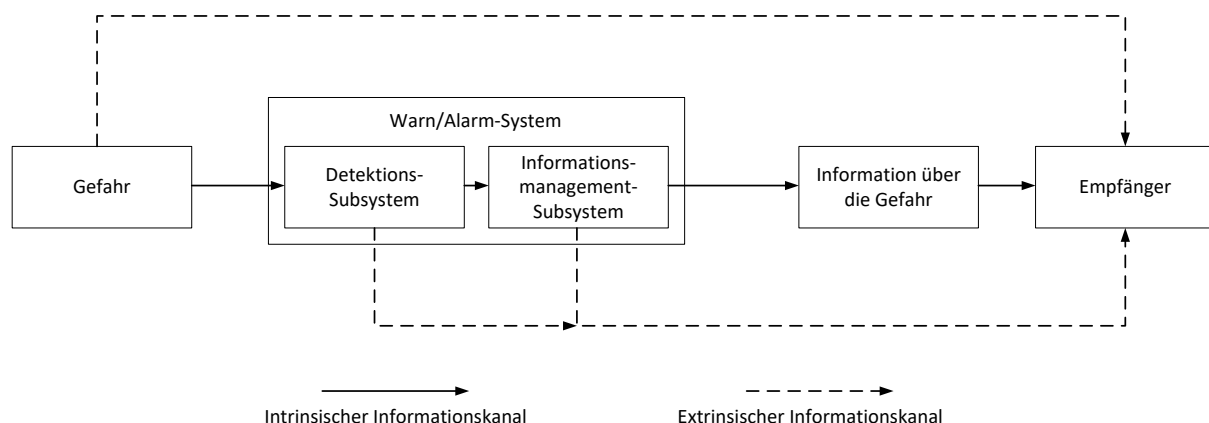
Bresnitz (1984) postuliert, dass der Cry Wolf Effekt ein Lernen aus Erfahrungen ist, das auf schließenden Prozessen basiert. Somit würde es unter die kognitiven Lerntheorien fallen und wäre nicht einfach nur ein Lernen im Sinne eines operanten

Konditionierens, in dem häufige Paarungen eines Alarmreizes mit einer antizipierten, aber ausbleibenden Konsequenz zu einer Missachtung des Alarmes führen.

Bresnitz berichtet von zwei vollkommen verschiedenen Ergebnissen des Schließens nach dem Erleben einer Episode mit falschen Alarmen, die sich jedoch im gleichen Verhalten äußern, nämlich einer schlechteren bzw. ausbleibenden Reaktion auf einen Alarm.

Das erste mögliche Ergebnis ist der bereits erwähnte Verlust des Vertrauens in das Alarmsystem, bzw. die niedrige Compliance in das Alarmsystem. Das zweite mögliche Ergebnis wäre eine Veränderung der wahrgenommenen Gefahr, vor der das Alarmsystem warnen soll. In diesem Fall bleibt die Compliance in das Alarmsystem hoch, jedoch wird das Ereignis vor dem gewarnt wird als nicht gefährlich erlebt. Dies kann insbesondere dann eine Gefahr darstellen, wenn die Gefahr vor der gewarnt wird, nicht offensichtlich, sondern eher abstrakt ist. Bezogen auf das Fahren mit einem hochautomatisierten Fahrzeug sind die Umweltgegebenheiten, die eine Fahrzeugautomation an ihre Systemgrenzen bringt, für den Fahrer nicht immer offensichtlich. Die Gefahr wird dem Fahrer erst dann bewusst, wenn es tatsächlich zu einem der Situation nicht angepassten Verhalten der Automation kommt. Eine Konsequenz, die aber nicht immer dann resultiert, wenn diese resultieren könnte.

Zu welchem Schluss eine Person kommt, hängt davon ab welche Informationen bzw. welche Beweise bzw. Hinweise für welche Art der Vermutung vorliegen. Bresnitz spricht hier von intrinsischer und extrinsischer Evidenz, bzw. dem extrinsischen und intrinsischen Informationskanal. Intrinsisch bedeutet in diesem Kontext, dass die Informationen über eine Gefahr aus dem Alarmsystem kommen. Extrinsisch bedeutet, dass die Informationen aus anderen Quellen als dem Alarmsystem selber kommen, so z.B. direkt aus der Umwelt oder aus anderen Systemen, neben dem Alarmsystem (siehe Abbildung 2).



**Abbildung 2: Extrinsischer und Intrinsischer Informationskanal nach Bresnitz (1984)**

### 2.9.1 Die Evidenztheorie des kausalen Lernens

Bresnitz spricht von Evidenzen für unterschiedliche Vermutungen bzw. Hypothesen des Empfängers, welche nach Art der Evidenz zu unterschiedlichen Schlüssen führen können.

Im Bereich der kognitiven Lerntheorien zum kausalen Lernen formulierte White (2000, 2002, 2003) die Theorie der Evidenzevaluation, welche gut auf die von Bresnitz aufgestellte Erklärung zur Entstehung eines Cry Wolf Effektes anzuwenden ist.

Die Theorie der Evidenzevaluation gehört zu den regelbasierten Theorien des kausalen Lernens. Nach Hagmayer & Waldmann (2006), welche eine gute Übersicht zu Theorien des kausalen Lernens erstellt haben, wird von jeder Person nach einer bestimmten Gesetzmäßigkeit bzw. einer Regel gesucht, mit der sich die An- oder Abwesenheit bestimmter Effekte nach einer beobachteten Ursache herleiten lässt. Diese Regel bzw. deren Stärke wird insbesondere über die Häufigkeit bzw. Regelmäßigkeit, mit der ein bestimmter Effekt nach einer Ursache beobachtet werden kann, oder eben nicht beobachtet werden kann, bestimmt. Dieser subjektiv statistische Zusammenhang zwischen Ursache und Wirkung bestimmt maßgeblich, ob eine Kausalrelation als vorfindlich oder nicht wahrgenommen wurde.

Nach Whites Theorie der Evidenzevaluation bilden Personen Hypothesen darüber wie Ursachen und Effekte miteinander zusammenhängen. Die beobachtete gemeinsame An- bzw. Abwesenheit von Effekten nach einer Ursache bilden Evidenzen für die Gültigkeit der Zusammenhangshypothese (konfirmatorische Evidenz). Das Ausbleiben der, aufgrund der Zusammenhangshypothese vermuteten, Effekte, stellt diskonfirmatorische Evidenz dar und kann die Stärke der Zusammenhangshypothese verringern.

Ein entsprechend hoher Anteil diskonfirmatorischer Evidenz kann so weit gehen, dass die Hypothese schließlich verworfen wird oder gar neue Hypothesen aufgestellt werden.

Nach White (2003) ist der Anteil der konfirmatorischen Evidenz von allen möglichen Konstellationen aus Paarungen der An- und Abwesenheit von Ursache und Effekt, maßgeblich für die Stärke der angenommenen Zusammenhangshypothese. Der Anteil der konfirmatorischen Evidenz ergibt sich aus der Vierfeldertafel (siehe Tabelle 2) und folgender Formel:

$$\text{Konfirmatorischer Evidenzanteil} = (A + D - B - C) / (A + D + B + C)$$



**Tabelle 2 – Vierfeldertafel mit den Kombinationen aus An- und Abwesenheit von Ursache und Effekt (nicht zu verwechseln mit dem ähnlich aussehenden Vierfelderschema der Signalentdeckungstheorie).**

	Effekt gegeben	Effekt abwesend
Ursache gegeben	A	B
Ursache abwesend	C	D

Je höher also der Anteil der Paarungen aus gemeinsamem Vorliegen oder gemeinsamer Abwesenheit von Ursache und Effekt, im Vergleich zu den Fällen, in denen nur entweder Ursache oder Effekt vorliegen, an der Gesamtmenge der erlebten Situationen, desto höher wird die Gültigkeit der Zusammenhangshypothese angesehen. Die Theorie erlaubt demnach Urteile darüber, ob überhaupt ein Kausalzusammenhang vorliegt, als auch über die Stärke des Kausalzusammenhangs (White, 2003, 2002).

Die Theorie erlaubt nach White (2003, 2002) ebenso Urteile über die Wahrscheinlichkeit, dass eine Ursache einen Effekt hervorruft. Diese Urteile über die Wahrscheinlichkeit, dass eine Ursache einen Effekt hervorruft, ist insbesondere interessant vor dem Hintergrund der Beobachtungen von Bliss, Gilson & Deaton (1995 – siehe Abschnitt Alarmzuverlässigkeit und Vertrauen in das Alarmsystem) welche in ihrer Studie beobachteten, dass etwa 90% der Versuchspersonen nicht immer auf unzuverlässige Alarme reagierten, sondern ihre Reaktionsrate von der erlebten Häufigkeit falscher Alarme abhängig machten, ein Verhalten das die Autoren „Probability Matching“ nennen. Die Autoren schließen, dass Versuchspersonen ihr Verhalten an die subjektiv wahrgenommene Wahrscheinlichkeit eines korrekten Alarms anpassen. Die Theorie der Evidenzevaluation passt recht gut auf dieses Verhalten, indem die Beobachtung von Ausbleiben von Merkmalen der Ursache (Automationsfehler bzw. Systemgrenze) nach dem Effekt (Alarm) zu einer Hypothese führen die keine eindeutige 1:1 Beziehung zwischen Ursache und Effekt postuliert. Anscheinend wird eher eine Hypothese gebildet, welche nur eine bestimmte Wahrscheinlichkeit über das Folgen eines Ereignisses nach einem Alarm zulässt.

Sicherlich spielt nicht nur die Anzahl bzw. relative Häufigkeit der gemeinsamen An- oder Abwesenheit von Alarm/Warnung und „Automationsfehler“ eine Rolle dabei, ob eine Hypothese verworfen wird, sondern auch eine Abwägung der Kosten, die es hätte, wenn die Hypothese verworfen wird, obwohl sie eigentlich gilt. Diese Kosten hängen davon ab, wie schwerwiegend die Konsequenz eingeschätzt wird, wenn auf einen Alarm einmal nicht reagiert wird. Das Risikoempfinden spielt also auch eine Rolle. Auch wenn die Eintrittswahrscheinlichkeit eines echten Automationsfehlers nach einer Warnung/Alarm als recht gering angesehen wird, der Zusammenhang

also eher als schwach eingeschätzt, das Eintreten der Konsequenz aber entsprechend negativ beurteilt wird, kann auch ein als sehr gering angenommener Zusammenhang zu einem recht stabilen Verhalten führen. (z.B. Roulston & Smith, 2004)

Bezogen auf den Cry Wolf Effekt lässt sich dessen Entstehung anhand der Theorie der Evidenzevaluation folgendermaßen erklären:

Eine Warnung oder eine Übernahmeaufforderung stellen einen Effekt einer für den Nutzer zunächst nicht eindeutig erkennbaren Ursache dar. Die meisten Ursachen für Automationsfehler, wie z.B. Sensorfehler durch Witterungseinflüsse, tiefstehende Sonne etc., sind nicht-Experten Nutzern zumindest zunächst nicht geläufig, weshalb eher ein implizites Systemverständnis in Form des Trusts in Automation gebildet wird. Damit eine Warnung als zutreffend bzw. gerechtfertigt empfunden wird, muss die Ursache dafür erkannt werden können. Das Vorhandensein der Ursache äußert sich für einen Automationsnutzer aber häufig nur „zuverlässig“ durch deren Folgen z.B. in Form eines beobachtbaren Automationsfehlers.

Solange die tatsächliche Ursache für die Warnung oder die Übernahmeaufforderung nicht erkannt werden kann (z.B. ein Zusammenhang zwischen Systemgrenzen, Warnungen und entsprechend ausbleibender oder nicht angemessener Automationsaktion), wird eher eine Art Kausalzusammenhang zwischen zwei Effekten der eigentlichen Ursache, nämlich der Warnung und dem danach beobachteten Automationsverhalten, hergestellt.

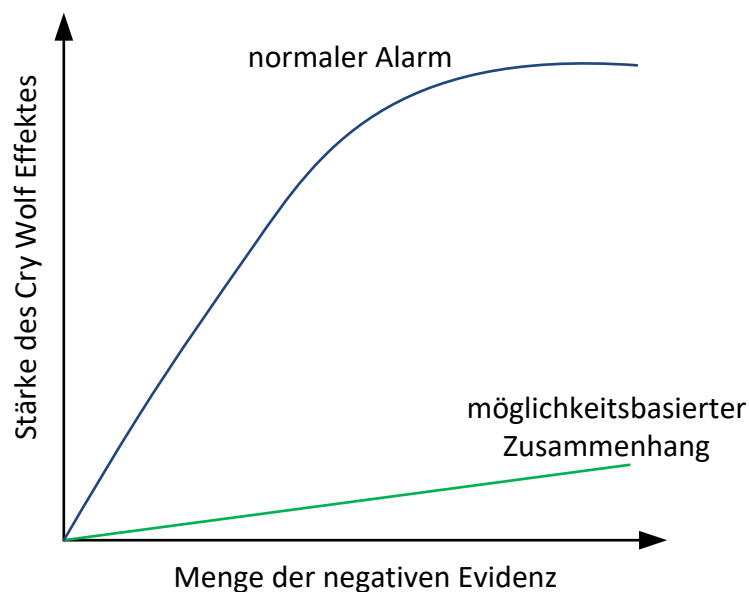
Die entsprechende Zusammenhangshypothese könnte am Anfang z.B. lauten: „Nach einer Warnung / Übernahmeaufforderung kommt es zu einem potenziell gefährlichen, weil der Situation nicht angemessenem, Automationsverhalten. Das Automationsverhalten nach einer Warnung oder einer Übernahmeaufforderung kann dann entweder als konfirmatorische oder diskonfirmatorische Evidenz zu dieser Hypothese gesehen werden. Bleibt nach einer Warnung das durch die Hypothese erwartete Ereignis mit einer bestehenden Regelmäßigkeit aus, sodass der konfirmatorische Evidenzanteil niedrig ist, kann die Hypothese entweder modifiziert werden, z.B. in der Art: „Nur auf wenige Warnungen folgt tatsächlich auch ein Automationsfehlverhalten“. Im schlimmsten Fall kann die Hypothese verworfen werden, in der Art „die Warnung hat keine prognostische Validität, sie erlaubt keine Aussage darüber, ob ein Automationsfehlverhalten folgt oder nicht“.

## *2.10 Möglichkeiten der Entstehung eines Cry Wolf Effektes zu entgegnen*

Abgeleitet von Whites Theorie der Evidenzevaluation müsste es möglich sein, die Cry Wolf Anfälligkeit von Alarm/Warnsystemen durch Beeinflussung der

Zusammenhangshypothese zu variieren. So könnte es z.B. möglich sein, bestimmte Zusammenhangshypothesen direkt zu vermitteln.

Diese Zusammenhangshypothese darf keinen streng monotonen Zusammenhang zwischen dem Vorliegen einer Warnung/Alarm oder Übernahmeaufforderung und einem der Situation unangemessenem Automationsverhalten in dem Sinne „wenn Alarm dann Konsequenz“ annehmen. Stattdessen muss eher die Möglichkeit eines Automationsfehlers in bestimmten Situationen aufgrund bestimmter Automationszustände bzw. Merkmale vermittelt werden (Abbildung 3).



**Abbildung 3: In Abhängigkeit der Alarmstrategie vermutete Beziehung der Menge der erlebten negativen Evidenz, auf die Neigung einen Cry Wolf Effekt zu entwickeln.**

Wichtig ist, dass das Nicht-Eintreten einer Konsequenz nach einem Alarm als eine genauso berechtigte Folge des Alarms erwogen wird, wie das Eintreten einer Konsequenz. Wesentlich ist hier, dass es möglichst auch nicht zu „Misses“ kommt. In den Situationen, in denen eben kein Alarm geäußert wird, darf auch keine „Konsequenz“ passieren.

Die in den meisten Untersuchungen zu Transitionen und Übernahmezeiten verwendete Warnungen sind „klassische“ Alarmstrategien die meist aus einer Paarung aus akustischem Hinweistönen und einem visuellen Element z.B. Text „Übernehmen“ oder Warnsymbol (z.B. Damböck, Farid, Töner & Bengler, 2012) oder einer akustischen Aufforderung „Take Over“ (z.B. Tofetti et al, 2009) bestehen. Melcher et al (2015) untersuchten auch andere Möglichkeiten der Kontrollrückgabe, wie z.B. einen Bremsruck oder eine Warnung die direkt in eine zu Fahrerablenkung genutzte Nebenaufgabe auf einem Smartphone dargeboten wurde. Obwohl keine dieser Untersuchungen die Erkundung einer möglichen Cry Wolf Effekt Anfälligkeit

der Rückmeldung zum Ziel hatte, fördern diese Arten der Rückmeldung wahrscheinlich das Bilden einer eher starren Zusammenhangshypothese in der Art: „Wenn Alarm, dann (Automations-)Fehler“. Sie liefern darüber hinaus auch keinerlei Erklärung warum es eine Warnung gab. Solche starren Zusammenhangshypothesen sind der Theorie nach eher anfällig für einen Cry Wolf Effekt.

In der Literatur finden sich zwei wesentliche Ansätze, die auf unterschiedliche Weise Einfluss auf die Bildung von Zusammenhangshypothese nehmen. Auch wenn sich diese Studien nicht explizit auf Whites Theorie der Evidenzevaluation stützen, so lässt sich die Theorie gut auf die jeweils genutzte Strategie anwenden. Die beiden Strategien sind

- Modifikation der Hypothese durch Vorwissen über die Automationszuverlässigkeit
- Modifikation der Hypothese durch die Art der Warnungs- / Alarmgestaltung.

#### 2.10.1 Einfluss von Vorwissen auf das Erleben falscher Alarme

Bliss und Dunn (1995) untersuchten eine Variante den Cry Wolf Effekt zu verringern, in welcher Versuchspersonen vor eigentlichem Versuchsantritt gesagt wurde, dass die Rate der falschen Alarme niedriger sei, als sie im Laufe des Versuchs tatsächlich war. Die Versuchspersonen bekamen die Information, der Alarm sei 75% zuverlässig, tatsächlich war dieser aber nur in 50% der Fälle zuverlässig.

Der Alarm selbst wurde auf einem Monitor durch das Wort „NOTE“ dargestellt und mit einem einfachen Glockenklang begleitet.

Diese Bedingung wurde mit jeweils 20 Versuchspersonen gegen eine Kontrollgruppe mit einem ebenfalls 50% zuverlässigen Alarm getestet. Im Versuch bearbeiteten die Versuchspersonen in einem Doppelaufgabenparadigma eine psychomotorische Aufgabe, während sie auf der anderen Seite auf das Alarmsystem reagieren sollten. Eine Alarmreaktion sollte durch einen Mausklick auf ein bestimmtes Feld auf dem Computermonitor erfolgen.

Es zeigte sich, dass falsche Information über die tatsächliche Zuverlässigkeit des Alarmsystems einen Effekt hatte, indem häufiger auf den Alarm reagiert wurde als in der Kontrollgruppe. Es wurde hier also ebenfalls ein Vorwissen erzeugt, dass es ein unzuverlässiges Alarmsystem geben wird, und, bezogen auf die Evidenztheorie, dadurch bereits eine „lockere“ Zusammenhangshypothese vermittelt. Diese führt dann jedoch, ausgehend von der „schwachen“ Zusammenhangshypothese zu einer Stärkung der Hypothese, wenn ein höherer Anteil konfirmatorischer Evidenz erlebt wird als angenommen. Eine solche Strategie ist aber wahrscheinlich nur am Anfang effektiv – im Laufe der zunehmenden Erfahrung mit dem System wird die Hypothese

demnach wieder angepasst, was wiederum zu einer gesteigerten Cry Wolf Neigung führen kann.

## **2.10.2      Auswirkung der Alarmgestaltung auf den Cry Wolf Effekt**

### **2.10.2.1      Einfluss der Alarmdringlichkeit**

In der gleichen Studie von Bliss und Dunn (1995) wurden auch Auswirkungen der Alarmgestaltung selbst untersucht. Hier wurde der Alarm mit unterschiedlichen Dringlichkeiten kommuniziert. In der Stufe niedrige Dringlichkeit wurde der Alarm genauso zurückgemeldet wie in der ersten Möglichkeit (Das Wort NOTE dargestellt auf weißem Hintergrund, begleitet von einem Glockenklang - siehe oben). In der Stufe hohe Dringlichkeit wurde das Wort „DANGER“ in Rot auf weißem Hintergrund auf dem Monitor dargestellt, das mit dem auditiven Signal eines Feuersalarms begleitet wurde. Auch hier war der Alarm nur in 50% der Fälle zuverlässig. Diese Bedingungen wurden wieder mit der Kontrollgruppe mit einem 50% zuverlässigen Alarm getestet (siehe oben).

Es zeigte sich, dass die Versuchspersonen in der Bedingung mit Dringlichkeitsinformation häufiger auf den Alarm reagierten als in der Kontrollgruppe. Jedoch zeigte sich nicht, wie von den Autoren erwartet, dass Versuchspersonen bei dringlichen Alarmen häufiger reagierten als bei weniger dringlichen. Sie reagierten tatsächlich ähnlich häufig. Es zeigte sich jedoch, dass Versuchspersonen bei hoher Dringlichkeit schneller reagierten. Die Kommunikation von Dringlichkeit scheint weniger die Zusammenhangshypothese selbst zu beeinflussen, sondern wohlmöglich eher die Wahrnehmung des Risikos. Wie im Abschnitt „Die Evidenztheorie des kausalen Lernens“ dargelegt, kann das empfundene Risiko nicht auf einen Alarm zu reagieren hoch sein, wenn die Eintrittswahrscheinlichkeit zwar als niedrig angesehen wird, eine mögliche Konsequenz aber als schwerwiegend angenommen wird.

### **2.10.2.2      Likelihoodalarne, Alarmcoding & Task-Critical Information**

Eine Rückmeldung die explizit die Unzuverlässigkeit der Warnung in der Gestaltung berücksichtigt, wurde unter anderem von Sorkin & Kantowitz (1988) untersucht. Sorkin & Kantowitz (1988) gingen der Frage nach, in wie weit die Wahrscheinlichkeit, die darüber entscheidet, ob ein Alarm geäußert wird oder nicht, explizit in die Gestaltung des Alarmes mit aufgenommen werden sollte. Sie führten neben den bisher üblichen Stufen kein Alarm und Alarm, Zwischenstufen, sogenannte Likelihoodstufen ein. Ein regulärer Alarm soll in dieser Art der Alarmgestaltung nur dann geäußert werden, wenn die Wahrscheinlichkeit, dass ein Grund für den Alarm auch tatsächlich vorliegt hoch ist. Likelihoodalarne werden geäußert, wenn die Wahrscheinlichkeit, dass kein Grund für den Alarm vorliegt zu niedrig ist und die

Wahrscheinlichkeit, dass ein Grund für einen tatsächlichen Alarm vorliegt auch zu niedrig ist. Im Sinne der Signalentdeckungstheorie kann das Antwortverhalten des Alarmsystems liberal sein (Tendenz in Zweifelsfall eher anzuschlagen), nur, dass es diesmal nicht direkt einen Alarm in dem Sinne äußert, sondern eine Wahrscheinlichkeit, eine „Likelihood“, für das Vorliegen eines Alarmes.

Auf diese Weise soll immer dann, wenn es auch einen berechtigten Alarmgrund gibt, eine Zuwendung zum Alarm stattfinden. Gleichzeitig soll in den niedrigeren Likelihoodstufen abgewogen werden, ob man dem Alarm näher nachgeht oder nicht. Dies würde ebenfalls die Leistung in einer Primäraufgabe steigern, da man sich von dieser weniger zum Zweck der Alarmbearbeitung abwenden muss.

Sorkin & Kantowitz untersuchten diese These mit einer recht kleinen Stichprobe von 6 Versuchspersonen. Die Versuchspersonen sollten in einem Doppelaufgabenparadigma eine Trackingaufgabe ausführen, in der mit einem Joystick ein Cursor einem sich bewegenden Objekt nachgeführt werden sollte. Ebenso wurde eine Monitoringaufgabe bearbeitet, in der anhand des Wertes dreier Zahlen geschlossen werden sollte, ob ein Signal vorliegt oder nicht. Dabei hieß es, je höher der Gesamtwert der Zahlen desto wahrscheinlicher ist das Vorliegen des Signals. Die Versuchspersonen bekamen nach jeder Entscheidung Feedback in der Form „Correct Rejection“, „False Alarm“, „Missed Signal“ & „Correct Alarm“. Die Versuchspersonen wurden in der Monitoring Aufgabe durch ein Alarmsystem unterstützt. Es wurden zwei Faktoren der Alarmgestaltung variiert. Einerseits das Alarmcoding in den Stufen verbal und Farbe und andererseits die Alarmstufen in den Stufen normaler Alarm (kein Alarm und Alarm) und Likelihood (kein Alarm, niedrige Likelihood, hohe Likelihood & Alarm).

Es zeigte sich, dass alle Arten der Rückmeldung eine deutliche Verbesserung im Trackingtask brachten, gegenüber keinem Alarm. Insgesamt stieg auch die Güte der korrekten Alarmresponses deutlich an. Es konnten keine Effekte des Alarmcodings gefunden werden. Ein tendenzieller Vorteil des vierstufigen Likelihoodalarms gegenüber dem normalen Alarm konnte nur in einer Bedingung eines schweren Trackingtasks beobachtet werden. Hier zeigte sich eine bessere Reaktion. In einer leichten Bedingung des Trackingtasks zeigte sich sogar eine schlechtere Alarmresponse in der Bedingung Likelihoodalarm als in der Bedingung normaler Alarm.

Die Autoren schließen, dass ein Alarmsystem mit mehr als zwei Stufen keine höhere mentale Belastung verursacht, durch ein Abwägen wann nun auf den Alarm reagiert werden sollte oder nicht. Aufgrund der geringen Stichprobengröße und des gewählten Designs der Messwiederholung sind die Aussagen eher zurückhaltend zu interpretieren.

Bustamante (2008) untersuchte im Kontext der Luftfahrzeugführung in zwei Simultorexperimenten die Effekte von Likelihood-Alarmstrategie und aufgabenspezifischer Zusatzinformation, jeweils unter der Variation von Workload, auf das Auftreten eines Cry Wolf Effektes.

Aufgabenspezifische Zusatzinformationen sollen neben dem anwesenden unreliaablen Alarmsystem den Piloten bzw. Versuchspersonen die Möglichkeit geben, die Berechtigung des Alarms anhand dieser Zusatzinformationen zu bestätigen oder eben abzuweisen. Dadurch soll nur noch auf berechtigte Alarme reagiert werden. Der von Bustamante vermutete Nachteil dieser Strategie ist, dass die Verabreichung der Zusatzinformation einen zusätzlichen Workload in einem teils eh schon fordernden Szenario erzeugt.

In der Likelihoodstrategie soll durch die Einführung der Likelihoodstufen den Versuchspersonen die Möglichkeit gegeben werden, besser zu entscheiden, wann sie sich einem Alarm zuwenden und wann eher nicht. So soll die Reaktion auf tatsächlich vorliegende Alarmgründe qualitativ und quantitativ erhöht werden und die Tendenz des „Probability Matchings“, der Anpassung der Reaktionswahrscheinlichkeit auf einen Alarm, anhand der wahrgenommenen Reliability des Alarms (Bliss, Gilson & Deaton, 1995) minimiert werden.

Bustamante nutzte zwei Experimente mit 30 und 24 Versuchspersonen, in denen er jeweils in einem Messwiederholdesign normale (binäre) und Likelihoodalarme, bei niedrigem und hohem Workload variierte. Der einzige Unterschied zwischen Experiment 1 und Experiment 2 war, dass in Experiment zwei zusätzlich aufgabenspezifische Zusatzinformationen gegeben wurden, anhand derer die Berechtigung eines Alarms leichter verifiziert werden konnte. In beiden Settings mussten die Versuchspersonen eine Trackingaufgabe sowie eine Ressourcen-Überwachungsaufgabe bearbeiten. Der Workload wurde dadurch variiert, dass in der Ressourcen Überwachungsaufgabe der hohen Workloadbedingung, häufiger Handlungen der Versuchspersonen, etwa durch Umpumpen von Treibstoff von einem in den anderen Tank, erforderlich waren.

Zusätzlich sollte eine Monitoringaufgabe ausgeführt werden in der die Triebwerke des simulierten Flugzeuges überwacht wurden, dabei wurden die Versuchspersonen durch ein Alarmsystem unterstützt. Das Alarmsystem hatte insgesamt eine niedrige Zuverlässigkeit von 18%. D.h. nur 18 von 100 Alarmen waren „echte“ Alarme. Beim binären Alarmdesign wurde der Alarm auf einem Display mit dem Wort „WARNING“ auf gelbem Hintergrund dargeboten, begleitet von einem Piep-Ton. Der Likelihoodalarm bei Bustamante hatte im Gegensatz zu den vier Stufen bei Sorkin & Kantowitz nur drei Stufen, kein Alarm, niedrige Likelihood (5% aller Alarme sind wahr) und hohe Likelihood (88% aller Alarme sind wahr). Der Alarm der niedrigen Likelihoodstufe war derselbe wie in der Bedingung binäres Alarmsystem. In der Stufe

hohe Likelihood wurde das Wort Alarm auf rotem Hintergrund begleitet von einem höheren Piep-Ton dargeboten.

Die Ergebnisse zeigen, dass die Likelihoodstrategie wie erwartet zu einer deutlich besseren Alarmreaktion führt. So wird deutlich häufiger auf tatsächliche Alarmer reagiert, welche alle in der Bedingung hohe Likelihood lagen. Die Autoren berichten, dass es im Fall der Likelihoodalarme zu deutlich besseren Entscheidungen führt, auf berechnigte Alarmer einzugehen und auf unberechnigte Alarmer eben nicht.

Kritik ist hier darin zu üben, dass eine sehr eindeutige Einteilung der beiden Likelihood Stufen in sehr reliabel (88%) und sehr wenig reliabel (5%) stattfand. Leider wird in der Studie nicht berichtet, wie häufig Versuchspersonen im Fall der niedrigen Likelihood auf einen berechtigten Alarm reagierten bzw. wie häufig dieser irrtümlich ignoriert wird. So verbessert z.B. die aufgabenspezifischen Zusatzinformationen die Leistung in der Entscheidung einem Alarm nachgehen zu wollen nur bei einem binären Alarm, brachte jedoch keine zusätzliche Verbesserung bei Likelihoodalarmen. Die Nutzung dieser Zusatzinformationen wäre aber zumindest in der niedrigen Likelihood Bedingung ein Indiz dafür gewesen, dass sich Nutzer hier eher mit der Verifizierung des Alarms beschäftigen.

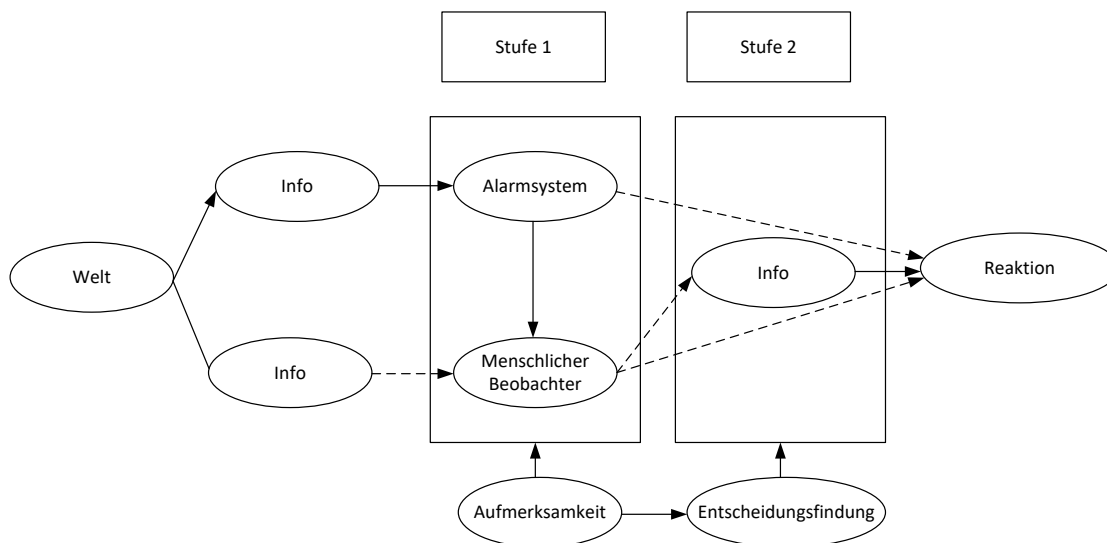
## *2.11 Einfluss von Likelihoodinformation auf die Hypothesenbildung*

Die genannten Studien zu Likelihoodalarmen liefern Hinweise, dass die Art der Alarmgestaltung einen Einfluss auf die Cry Wolf Neigung haben kann. Bezogen auf die Evidenztheorie des kausalen Lernens kann es sein, dass durch die Abstufung des Alarms in zwei (oder mehr) Stufen, je Stufe eine eigenständige Hypothese gebildet wird. So kann z.B. eine Hypothese bezüglich der Likelihoodinformation – in der ein eher schwacher Zusammenhang vermutet wird und eine weitere Hypothese die einen recht starken Zusammenhang zwischen regulärer Warnung und einem Automationsfehler postuliert. Die Likelihoodstufe sollte zu einer Hypothese führen in der Explizit eine Möglichkeit eines Automationsfehlers nach einem Alarm erwogen wird, jedoch nicht notwendigerweise das sichere Auftreten eines solchen Automationsfehlers.

Likelihoodalarme wurden in erster Linie in Kontexten untersucht – in denen üblicherweise eine Verifizierung stattfindet, ob ein Alarm berechnigt ist oder nicht. In Bereichen der Überwachung komplexer Systeme wie z.B. Luftfahrzeuge (z.B. Bustamante, 2008) oder Leitwarten (z.B. Sorkin & Kantowitz, 1988) wird nach einem Alarm immer erst untersucht, um was für einen Alarm es sich handelt und ob dieser berechnigt ist.



Sorking & Kantowitz (1988) als auch Bustamante (2008) ziehen hier ein Zweistufenmodell der Informationsverarbeitung heran. In diesem Modell wird in einer ersten Stufe der Alarm zur Kenntnis genommen und dadurch entsprechend Aufmerksamkeit auf ein (mögliches) Problem gelenkt, ebenso wird er einer ersten Beurteilung unterzogen. In dieser ersten Stufe wird beurteilt, ob es sich um einen eher zuverlässigen oder einen eher unzuverlässigen Alarm handelt. Der Ausgang dieser ersten Beurteilung beeinflusst, wie weiter in der zweiten Stufe verfahren wird. Ob und wieviele zusätzliche Informationen eingeholt werden müssen, um einen Alarm zu bestätigen oder eben ignorieren zu können (siehe Abbildung 4).



**Abbildung 4: 2 Stufiges Signal-Entdeckungsmodell der Entscheidungsfindung nach Bustamante (2008) – Gestrichelte Linien stellen optionale Beziehungen dar welche von den Eigenschaften einer jeweiligen Situation abhängen**

Dieses Verifizieren kann im Fall von unzuverlässigen Alarmen recht viel Zeit in Anspruch nehmen und entsprechend einen Systembediener von seiner eigentlichen Aufgabe ablenken. Ein Likelihoodalarm soll aufgrund der Alarmstaffelung dazu beitragen, von vornherein besser unterscheidbar zu machen, für welche Arten von Alarmen noch eine mehr oder weniger umfangreiche Verifizierung notwendig ist (Likelihoodalarm) und welchen Alarmen eigentlich soweit vertraut werden kann, dass hier keine umfangreiche Verifizierung mehr durchgeführt werden muss. So soll insgesamt die Zuverlässigkeit mit der auf das Alarmsystem reagiert wird gesteigert werden und andererseits der Workload verringert werden, indem sich nur noch mit der Verifizierung der Alarme mit niedriger Likelihood gekümmert wird. Überträgt man die Likelihood-Alarmstrategie auf teil-, bzw. hochautomatisierte Fahrzeuge, so ist es hier nicht so, dass der Fahrer durch einen Alarm von seiner eigentlichen Hauptaufgabe abgelenkt wird. Im Zweifelsfall ist die Übernahme der Kontrolle von einer Automation und die Ausführung der Fahraufgabe immer die

Hauptaufgabe. Es ist hier sogar erstrebenswert, dass sich der Fahrer bei einem Alarm vermehrt mit der Überprüfung dieses Alarms beschäftigt. Diese vermehrte Zuwendung zu einer Aufgabe zwecks Überprüfung sollte ja durch die Stufe niedrige Likelihood gefördert werden. Im Auto würde sich so z.B. ein Fahrer vom hochautomatisierten Fahren in das teilautomatisierte Fahren begeben, der Fahraufgabe also mehr Aufmerksamkeit schenken, um im Fall eines berechtigten Alarms gut vorbereitet zu sein.

In den Studien von Sorkin & Kantowitz und insbesondere bei Bustamante wurde ein mögliches Problem des Likelihoodalarms klar. Wird bei den Stufen mit hoher Likelihood noch gut auf den Alarm reagiert, wird hingegen in der niedrigen Likelihoodstufe – der Cry Wolf Effekt auf diese Stufe konzentriert. Die Alarmstufe mit niedriger Likelihood wird so gesehen zu einer Art „Opfer-Alarm“, den man zugunsten der Alarme mit hoher Likelihood opfert. Eine weitere These hierzu liefern Clark, Ingebritsen & Bustamante (2010) welche durch die Einteilung der Likelihoodalarme in Stufen die Möglichkeit sehen, dass Nutzer solcher Alarme nicht nur ein „uncertainty mapping“ auf die bestimmten Stufen vornehmen, sondern auch ein „urgency mapping“ d.h. Alarme die nicht eine hohe Likelihood darstellen, werden als weniger dringend empfunden weil die Information über die unterschiedliche Likelihood (z.B. OK, Warning, Alarm) auch als Information bezüglich der Dringlichkeit interpretiert wird. Konsequenz durch beide Möglichkeiten ist, dass eine vermehrte Überprüfung und Verifizierung der Alarme mit niedriger Likelihood gar nicht, oder nur sehr wenig stattfindet. Dies ist in solchen Systemen in denen recht zuverlässig eine Grenzlinie gezogen werden kann, wann ein Alarm wahrscheinlich ist und wann nicht von Vorteil. Ist es aber nicht ohne Weiteres möglich diese Grenze zu ziehen – so steigt die Wahrscheinlichkeit, dass es auch in der Likelihoodstufe berechnete Alarme gibt. Durch die stattfindende Cry Wolf-Inflation auf der Stufe steigt die Wahrscheinlichkeit, dass auf berechnete Alarme unter den Likelihoodalarmen nicht mehr reagiert wird.

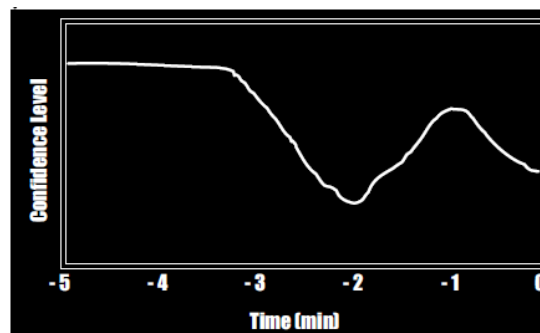
### 2.11.1 Likelihoodinformation durch Rückmeldung von Automationsunsicherheit

Eine weitere Möglichkeit die Unzuverlässigkeit von Alarmsystemen zu kommunizieren besteht darin, keine Unterteilung mehr in Likelihoodstufen vorzunehmen und dadurch die Aufteilung in „guter Alarm“ und „schlechter Alarm“ zu vermeiden, sondern ausschließlich eine einzelne Likelihoodartige Stufe zurückzumelden. Dieser Ansatz nutzt die Rückmeldung von Unsicherheitsinformation.

Bei einer Unsicherheitsrückmeldung liegt kein strenger Zusammenhang zwischen Rückmeldung und tatsächlichem Automationsverhalten vor. Verschiedene Ausgänge der Situation sind per se möglich. Wenn eine Zusammenhangshypothese in dieser

Art gebildet wird, müsste diese relativ unanfällig gegenüber dem Cry Wolf Effekt sein. So gibt es in der Literatur mehrere Ansätze effektiv Unsicherheit zurückzumelden – in denen es eine gewisse Wahrscheinlichkeit für falsche Alarme gibt. Beispiele sind die Unsicherheitsrückmeldung in der Bilderkennung medizinischer bildgebender Verfahren (z.B. Grigoryan & Rheingans, 2002) oder in der Identifizierung „Freund oder Feind“ im Kontext militärischer Flugzeuge (z.B. Banbury, Selcon, Endsley, Gorton & Tatlock 1998) sowie vielen Weiteren. Hier soll exemplarisch auf zwei Studien näher eingegangen werden.

Sarter & McGuirl (2006) untersuchten die Auswirkungen der Rückmeldung von System (Un)Sicherheit in einer Entscheidungssituation, in der es darum ging, die Zuverlässigkeit eines Vereisungs-Warnsystems im Flugzeug einzuschätzen. 30 Versuchspersonen, allesamt Piloten, sollten in einem Bewegungsimulator, der in der Lage war für Vereisung typische Bewegungs-Hinweise abzubilden, je 28-achtminütige Flüge durchführen. Dabei sollten sie einen durch Vereisung bedingten Strömungsabriss durch entsprechende Manöver vermeiden. Die Piloten wurden von einem System unterstützt, das Vereisung erkennen konnte und darauf bezogen Handlungshinweise zur Vermeidung eines Strömungsabrisse gab. Die Anweisungen des Systems waren aber im Mittel nur zu 70% zuverlässig. Zu diesem Vereisungs-Assistenten gab es ein Display, das die aktuelle Zuverlässigkeit des Assistenten, als auch den Verlauf der Zuverlässigkeit in den letzten 5 Minuten darstellte (Abbildung 5).



**Abbildung 5: Confidence Display aus der Studie von McGuirl & Sarter (2006)**

Es wurde ein Doppelaufgabenparadigma genutzt. Während der simulierten Flüge sollten die Piloten eine Nebenaufgabe bearbeiten, indem die Flugbewegungen anderer Flugzeuge anhand eines „Airtraffic Control Displays“ beobachtet und berichtet werden mussten.

Die Piloten wurden einer von zwei Bedingung zugeordnet. In der ersten Bedingung wurden feste System-(Un)Sicherheit Informationen zurückgemeldet, in der zweiten Bedingung kontinuierlich aktualisierte (Un)Sicherheitsinformation über die Zuverlässigkeit des Vereisungswarnsystems zurückgemeldet. In der Bedingung

aktualisierte Information wurden allen Piloten über die Flüge hinweg 9 Situationen mit stetig hoher Systemsicherheit (89% Reliabilität), 10 Situationen mit variabler System (Un)Sicherheit (durchschnittlich etwa 50%) und 9 Situationen mit niedriger System- (Un)Sicherheit (etwa 25%) ausgesetzt. Im Durchschnitt ergab dies eine Systemsicherheit von 70%. In der Bedingung feste (Un)Sicherheitsinformation wurde den Piloten permanent die gleiche Systemzuverlässigkeit von 70% zurückgemeldet. Es zeigte sich, dass es in der Bedingung feste System (Un)Sicherheitsrückmeldung etwa doppelt so häufig zu einem Strömungsabriss kam (etwa 64% aller Fälle) als in der Bedingung der kontinuierlich aktualisierende (Un)Sicherheitsinformation (etwa 36% aller Fälle). Es zeigte sich auch, dass die Piloten der „festen“ Bedingung deutlich häufiger den Anweisungen des Systems folgten, als es durch die angegebene System-(Un)Sicherheit gerechtfertigt gewesen wäre. In der Studie gab es keine Bedingung in der keine Systemunsicherheit zurückgemeldet wurde. So ist leider keine Aussage möglich, ob auch die Angabe der festen 70% System Sicherheit, einen Vorteil gegenüber gar keiner System-(Un)Sicherheitsrückmeldung ergeben hätte.

In einem anderen Kontext untersuchten Antifakos, Kern, Schiele & Schwaninger (2005) in wie weit sich die Rückmeldung der System-(Un)sicherheit bei „Context aware mobile phones“ auf das Vertrauen der Nutzer in diese Systeme auswirkte. Context aware bedeutet in diesem Zusammenhang, dass das Telefon anhand bestimmter situativer Merkmale erkennen kann, in welchem situativen Setting sich der Nutzer gerade befindet und dementsprechend automatisch anpasst, wie auf einen Anruf oder eine Mitteilung aufmerksam gemacht werden soll (gar nicht, Vibrieren, leises Klingeln, lautes Klingeln). 14 Versuchspersonen wurden Videos unterschiedlicher Situationen gezeigt. Die unterschiedliche Kritikalität dieser Videos bezüglich der Art wie auf Anrufe etc. hingewiesen wird, wurde in einer Voruntersuchung mit den gleichen Versuchspersonen bestimmt. Hier sollten diese angeben, wie sie in einem jeweiligen Szenario notifiziert werden möchten. Ebenso sollte angegeben werden wie wichtig es für den Nutzer selbst oder für seine Umgebung ist, dass er der Situation angemessen notifiziert wird. Im Experiment wurde untersucht, in wie weit sich die Rückmeldung von (Un)Sicherheitsinformation auf das Vertrauen auswirkt, die der Situation angemessenen Notifizierungsmodalität ausgewählt zu haben (gar keine, Vibrieren, leises Klingeln, lautes Klingeln). Vertrauen wurde dadurch erfasst, in wie weit die Systemeinstellungen in einer jeweiligen Situation überprüft wurden (kein Trust) oder eben nicht überprüft wurden (Trust). Im Experiment wurde innerhalb der Versuchspersonen über Videos hinweg die Art der Situation (kritisch, weniger kritisch, unkritisch) das rückgemeldete System(un)sicherheitslevel (90% sicher und

44

50% sicher) und die Verfügbarkeit von Kontexthinweisen für die automatische Entscheidung des Systems variiert.

Es zeigte sich, dass das Vertrauen in das System in kritischen Situationen am niedrigsten ist. Der (Un)Sicherheits-Level hat jedoch einen Effekt auf das Vertrauen der Nutzer. So wurden in der Bedingung hohe System Sicherheit tendenziell seltener die Systemeinstellungen überprüft als in der Bedingung niedrige Systemsicherheit. Wie auch in der vorangegangenen dargestellten Studie, wurde hier nicht untersucht, wie sich die Rückmeldung keiner (Un)Sicherheitsrückmeldung auf das Entscheidungsverhalten der Nutzer ausgewirkt hat. Es wird aber deutlich, dass über das Wissen des Nutzers über die zurückgemeldete System-(un)sicherheit, situativ das Vertrauen an ein System von variabler Zuverlässigkeit angepasst werden kann.









### 2.11.2 Likelihoodinformation bei assistiertem & hochautomatisierten Fahren

Im Kontext des assistierten bis hochautomatisierten Fahrens liegen mittlerweile einige Forschungen zu unterschiedlichen Alarmstrategien & Alarmgestaltungen vor, welche explizit eine hohe Variabilität in der Unsicherheit der vom Automationssystem erfassten Umgebungsinformationen berücksichtigt und in die Art einer Fahrerinformation integriert wird. Die im Folgenden dargestellten Studien untersuchten vor allem, wie Unsicherheitsinformation von Fahrern genutzt wird, bzw. ob die Darbietung von Unsicherheitsinformation überhaupt effektiv sein kann. Es wurde in diesen Studien nicht untersucht in wie weit Unsicherheitsinformation dazu beitragen kann, die Entstehung eines Cry Wolf Effektes zu vermeiden. Ein interessanter Aspekt der Unsicherheitsrückmeldung der in einigen der Untersuchungen deutlich wird ist, dass die Rückmeldung von Unsicherheit nicht nur als eine Art Warnung funktioniert, sondern auch direkt als Erklärung für ein nicht angemessenes Automationsverhalten. Die Warnung wird damit nicht nur ein weiterer Effekt einer verborgenen Ursache – sondern kommuniziert die Ursache für einen möglichen Automationsfehler direkt mit, nämlich Automationsunsicherheit. Allerdings muss hier erwähnt werden, dass die Ursache für diese Automationsunsicherheit auch nicht klar wird.

Duschl, Just & Niedermaier (2011) untersuchten einen Ansatz, der es dem Fahrer ermöglichen sollte mit den unsicheren Informationen, auf welche sich ein Verzögerungsassistent stützt, umzugehen. Der Verzögerungsassistent konnte Verkehrsschilder mit einer Geschwindigkeitsbegrenzung erkennen. Die Erkennung war dabei nicht immer eindeutig. Es wurden vier unterschiedliche Varianten der Verkehrszeichenerkennung verglichen. In drei dieser Rückmeldungen wurde

Unsicherheit in der Erkennung mal mehr mal weniger explizit kommuniziert. Die unterschiedenen Rückmeldungen waren: (siehe auch Abbildung 6)

- Optimistische Rückmeldung: Unsichere Daten wurden wie sichere Daten behandelt, was am Ehesten mit herkömmlichen binären Alarmen zu vergleichen ist.
- Pessimistische Rückmeldung: Unsichere Daten wurden gar nicht gezeigt, nur korrekte Daten wurden gezeigt.
- Vorsichtige (Cautious) Information: Unsichere Information wurde als solche gekennzeichnet
- Sehr vorsichtige (very cautious) Information: Unsichere Information wurde als solche gekennzeichnet und sichere Information wurde explizit als sicher gekennzeichnet.

	Uncertain Data	Certain Data
Optimistic		
Pessimistic		
Cautious		
Very Cautious		

**Abbildung 6: Rückmeldungsvarianten der Verkehrszeichenerkennung (aus Duschl, Just & Niedermaier, 2011)**

Der Fokus der Untersuchung lag hier nicht in der Frage, ob ein Cry Wolf Effekt durch eine entsprechende Alarmgestaltung vermieden werden kann, sondern eher auf der subjektiven Einschätzung der Rückmeldung hinsichtlich empfundenem Vertrauen in das System, der empfundenen Sicherheit, ob das System als funktionierend oder nicht wahrgenommen wird und ob das System insgesamt gemocht wird.

Die Erhebung wurde als Messwiederholungsdesign mit 30 Benutzern in einem Fahrsimulator durchgeführt. Die Ergebnisse zeigen, dass die optimistische Information subjektiv in fast allen erhobenen Dimensionen, insbesondere in den Punkten Vertrauen in das System und wahrgenommene Fehlfunktion, deutlich schlechter bewertet wurde, als die alternativen Rückmeldungen, in denen der unsichere Zustand kommuniziert wurde. Die Rückmeldungen in denen Unsicherheit explizit kommuniziert wurde (cautious & very cautious), wurden im Schnitt in allen

untersuchten Situationen mit Unsicherheit als hilfreicher eingeschätzt als das optimistische Design und das pessimistische Design (mit indirekter Unsicherheitsrückmeldung - unsichere Daten werden nicht gezeigt).

Die Autoren schlussfolgern, dass die Darbietung von Unsicherheitsinformationen im Kontext von Assistenzsystemen durchaus akzeptiert und von den Versuchspersonen als Mehrwert wahrgenommen wird.

Seppelt (2009) untersuchte eine Strategie der kontinuierlichen Rückmeldung der Systemgrenzen eines ACC Systems zur Optimierung des Umgangs eines Fahrers mit imperfekter Automation durch Förderung eines angemessenen Systemvertrauens.

Neben anderen Möglichkeiten den Fahrer kontinuierlich darüber zu informieren wie nah ein adaptiver Abstandregeltempomat an möglichen Systemgrenzen ist, untersuchte Seppelt auch die Möglichkeit Automationsunsicherheit als kontinuierliche Information bezüglich der Fähigkeit die ACC Unterstützung aufrechtzuerhalten zurückzumelden. Die Untersuchung fand mit 48 Versuchspersonen in einem statischen Fahrsimulator statt.

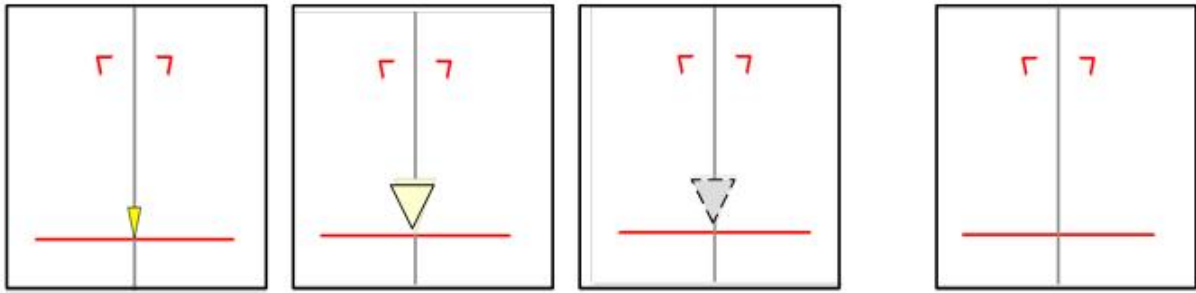
Das System war unsicher in Situationen mangelhafter Datenlage (Sensor Degradation) verursacht z.B. durch Regen unterschiedlicher Stärke. Seppelt verglich hier zwei Möglichkeiten der Rückmeldung von Unsicherheit.

- Visuell: Abnahme der Farbintensität. Ein abschwächen der normalen Displayfarbe von einem intensiven Gelb zu einem Hellgelb bis hin zu gar keiner Darstellung des Elementes im Fall eines Sensorausfalls (Abbildung 7) & Merkbare Unregelmäßigkeiten in der Displayaktualisierung
- Auditiv: Vermittlung von Unsicherheit durch ein zunehmend verzerrtes Audiosignal

Die Stärke der kommunizierten Unsicherheit hing dabei von der Stärke des Regens ab (siehe Tabelle 3 & Abbildung 7).

**Tabelle 3: Audiosignale und visuelle Charakteristika der Unsicherheitsrückmeldung in Abhängigkeit der Regenstärke (aus Seppelt, 1995)**

Regenstärke	Audioverzerrung	Füllung Farbe (RGB)	Füllung - Muster	Umrandung Farbe (RGB)	Linienstärke	Linien-Muster
<0,2	keine	255,255,0	massiv	0,0,0	3	durchgehend
0,2 - <0,4	gering	250,250,250	massiv	180,180,180	2	durchgehend
0,4 - < 0,6	mittel	240,249,240	massiv	220,220,220	1	unterbrochen
>=0,6	hoch	200,200,200	gestreift	150,150,150	1	unterbrochen



**Abbildung 7: Abnehmende Intensität der Farbe eines Displayelementes von einem intensiven Gelb über weiß, bzw. grau bis hin zu einem vollkommenen Fehlen des Elementes in einem Display zur kontinuierlichen Darstellung von ACC Systemgrenzen, soll eine abnehmende Fähigkeit eines Sensors die Umwelt akkurat zu erfassen, darstellen (aus Seppelt, 2009)**

Die Versuchspersonen sollten in einem Doppelaufgabenparadigma mit einem Fahrzeug über eine Strecke fahren und dabei zum einen die Querführung übernehmen, während sie in der Längsführung durch ein ACC System unterstützt wurden. Während der Fahrt fuhr das Fahrzeug einem Vorderfahrzeug nach, das in unregelmäßigen Abständen seine Geschwindigkeit variierte. In 50% der Fälle kam es, aufgrund unterschiedlicher Einflussfaktoren, zu einem nicht angemessenen Verhalten der Automation. In diesen Situationen mussten die Versuchspersonen eingreifen und z.B. bremsen. In der Bedingung, in der die Unsicherheitsrückmeldung untersucht wurde, führte unterschiedlich starker Regen zu unterschiedlich hohen Ausfallwahrscheinlichkeiten. Analog zu der Intensität des Regens wurde die Rückmeldung auf dem Display bzw. auditiv variiert (siehe Tabelle 3).

Zusätzlich zu dieser Fahraufgabe sollten die Versuchspersonen eine Zweitaufgabe bearbeiten. Diese bestand darin, auf Werbetafeln links und rechts der Straße zu achten und bei einem bestimmten Inhalt der Tafeln einen Knopf zu drücken. Die Nebenaufgabe diente dazu Ablenkung von der Fahraufgabe zu provozieren, bzw. auch Effekte der Rückmeldung in Form von unterschiedlich starker Hinwendung zur Nebenaufgabe, zu messen.

Die Versuchspersonen wurden instruiert, dass sie sich auf das ACC verlassen sollten, wenn es die Fahrsituation bewältigen kann – sie wurden auch instruiert das ACC System zu überwachen und es dann auszuschalten sobald sie dies anhand des System-Feedback für nötig hielten und es erst dann wieder einzuschalten sobald sie sich wieder sicher waren.

Die Faktoren Display type (kontinuierlich vs. diskret wobei diskret) und Modalität (visuelle Rückmeldung vs. auditive Rückmeldung) wurde zwischen den Gruppen variiert (siehe Tabelle 3). Innerhalb der Gruppen wurden die Fahrsituation bzw. die Art der erlebten Systemgrenzen variiert.

Es zeigte sich insgesamt ein deutlich positiver Effekt der kontinuierlichen Rückmeldung von Systemgrenzen verglichen mit der diskreten Rückmeldung. Es

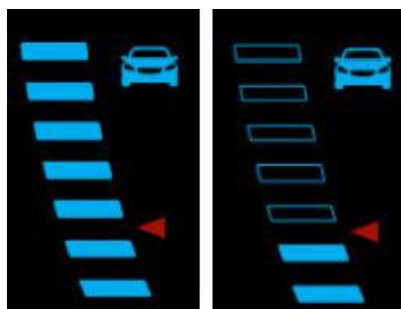


konnte allerdings kein nennenswerter Effekt der kontinuierlichen Rückmeldung der Automationsunsicherheit in den Situationen mit Regenfällen beobachtet werden. Mögliche Gründe hierfür liegen darin, dass die Bedeutung der Rückmeldung nicht instruiert wurde. Die Rückmeldung der Unsicherheit basierte auf einer Strategie die von Finger & Bisantz (2002) zwar in einem anderen Kontext, als recht effektiv identifiziert wurde, jedoch war den Versuchspersonen anscheinend nicht klar welche Bedeutung die Variation des Displays in diesem Kontext wohl haben könnte. Die Untersuchungen von Seppelt hatten auch nicht zum Ziel eine Rückmeldestrategie zu finden, welche die Entstehung eines Cry Wolf Effektes mindern sollte. Es sollte vielmehr untersucht werden, ob die Kommunikation des Sachverhalts unsicherer Automation an den Fahrer, generell eine effektive Methode darstellt dem Fahrer ein angepasstes Vertrauen in das Automationssystem zu ermöglichen und dadurch letztendlich bessere Kontrolle in Übernahmesituationen.

### 2.11.3 Rückmeldung einer Begründung für Kontrollrückgaben

Eine Unsicherheitsrückmeldung kann, neben der alarmierenden Komponente, direkt als Begründung für mögliche Automationsfehler funktionieren. Ein weiterer Vorteil neben der potenziell geringen Anfälligkeit für einen Cry Wolf Effekt wäre eine bessere Anpassung des Vertrauens in die Automation dadurch, dass Nutzern der Automation direkt vermittelt wird, dass die Automation eben situativ unsicher sein kann.

Helldin, Falkman, Riveiro & Davidsson (2013) untersuchten in einer Simulatorstudie mit 59 Versuchspersonen, in wie weit sich die kontinuierliche Rückmeldung der Automations-(un)sicherheit auf die Übernahmezeit, das Vertrauen der Fahrer in ein hochautomatisiertes Fahrzeug, als auch die Beschäftigung der Fahrer mit Nebenaufgaben auswirkt. Die Versuchspersonen wurden in zwei Gruppen eingeteilt, eine dieser Gruppen erhielt die Unsicherheitsinformation (Abbildung 8), die andere Gruppe erhielt keine Information.



**Abbildung 8: kontinuierliche Rückmeldung der Automations(un)sicherheit aus der Studie von Helldin et al (2013). Der rote Pfeil markiert die Grenze, an der die Automationsunterstützung nicht mehr gewährleistet werden kann.**

Die etwa 9-minütigen Fahrten fanden auf einer Landstraße unter variierenden Wetterbedingungen statt. Die Versuchspersonen wurden instruiert, dass die Wetterbedingungen Auswirkungen auf die Fähigkeit des Automationssystems die Unterstützung aufrechtzuerhalten, haben. Sie wurden ebenso instruiert, dass es ihnen überlassen sei, zu entscheiden, wann sie die Kontrolle über das Fahrzeug übernehmen, bzw. wann es nicht mehr angemessen erschien das Fahrzeug weiter kontrollieren zu lassen. Während der Fahrt kam es zu variierenden Schneefallbedingungen. Bei starkem Schneefall konnte es passieren, dass die Automation die Unterstützung einstellte – dies geschah indem das Auto plötzlich vom Gas ging, aber auch in Kurven nicht mehr lenkte. In den Szenarios fand dies immer in einer leichten Kurve statt, sodass Versuchspersonen hier lenken und optional auch bremsen mussten. Nach jeder Fahrt wurde das Vertrauen in das System anhand eines Fragebogens erfasst.

Im Mittel konnten die Versuchspersonen der Unsicherheitsbedingung nach etwa 1,9 Sekunden die Kontrolle über das Fahrzeug übernehmen, die Gruppe „keine Information“ brauchte hingegen 3,2 Sekunden. Der Unterschied war signifikant. Die Versuchspersonen der Unsicherheitsgruppe verbrachten allerdings mehr Zeit damit ihre Aufmerksamkeit von der eigentlichen Fahraufgabe abzuwenden (etwa 18% der Zeit vs. 12% der Zeit in der Bedingung ohne Information, Effekt ebenfalls signifikant). Bezüglich des subjektiv empfundenen Vertrauens wurde das System ohne Unsicherheitsrückmeldung als vertrauensvoller empfunden als das System mit Unsicherheitsrückmeldung. Die Autoren schließen, dass die Unsicherheitsrückmeldung zu einer Anpassung des Vertrauens zu einem Ausmaß der die tatsächliche Leistungsfähigkeit des Systems widerspiegelt, unterstützt. Ebenso verließen sich die Versuchspersonen anscheinend auf die Rückmeldung und lenkten sich entsprechend ab, wenn das Display eine genügend hohe Sicherheit anzeigte bzw. waren besser vorbereitet auf eine mögliche Kontrollrücknahme, wenn die Sicherheit unter ein bestimmtes Niveau fiel. Die Fahrer der Bedingung ohne Rückmeldung hatten hingegen keine eindeutige Grenze, wann die Kontrolle übernommen werden sollte, bzw. ab wann es sich lohnt sich mehr mit der Fahraufgabe zu beschäftigen.

Erstaunlicherweise führte in der Bedingung ohne Rückmeldung die eigentlich höhere Unsicherheit auf Seiten der Fahrer, wann es zu einem Automationsausfall kommen kann, zu einem höheren Vertrauen in die Automation als in der Bedingung in der die transparente Rückmeldung der Automationsleistungsfähigkeit stattfindet. Die Untersuchung zeigt auch, dass eine unscharfe Unsicherheitsrückmeldung, die keine eindeutige Aussage über das Auftreten einer bestimmten Konsequenz ermöglicht, effektiv sein kann, obwohl Unsicherheit nicht bedeutet, dass ein Automationsausfall tatsächlich auch passiert. In der Studie wurde allerdings nicht untersucht, in wie weit

häufiges Ausbleiben eines Automationsausfalls nach einer Unsicherheitsrückmeldung – im weiteren Verlauf zu einer stabilen Reaktion auf eine solche Rückmeldung führen kann. Es kam nur gegen Ende der Fahrt zu einer Rückmeldung hoher Automationsunsicherheit, der dann auch direkt der Automationsausfall folgte. Aussagen bezüglich einer möglichen Cry Wolf Resistenz von Unsicherheitsinformation sind also nicht möglich.

Die Effektivität einer Begründung von Automationsfehlern hinsichtlich der Handhabung von Automationsfehlern und dem Vertrauen in die Automation konnte auch z.B. von Bagheri und Jamieson (2004) aufgezeigt werden.

Bagheri & Jamison (2004) untersuchten in zwei Studien wie sich Informationen darüber wie zuverlässig die Automation in bestimmten situativen Kontexten ist, auf die Detektionsleistung von Automationsfehlern während einer Flugsimulation auswirkten (Studie 2). Sie lieferten damit auch eine vorab Begründung, warum in manchen Situationen mit einer schlechteren Systemleistung zu rechnen sei. Sie vergleichen diese Bedingung mit einer Bedingung, in der die Versuchspersonen diese Information über die kontextbezogene Zuverlässigkeit der Automation nicht bekamen. Es zeigte sich, dass vor dem simulierten Flug vermitteltes Wissen über die kontextabhängige Zuverlässigkeit der Automation zu einer deutlich besseren Detektionsleistung von Automationsfehlern führte. Die Autoren begründeten die Wirkung der Information über die kontextbezogene Zuverlässigkeit damit, dass die Versuchspersonen die Automation und ihre Funktion nun besser verstanden und daher problematische Situationen besser erkennen und antizipieren konnten, wodurch sie dann ihre Aufmerksamkeit situationsbedingt anders verteilten. Durch die Erhöhung des expliziten Wissens über die Automation, konnte demnach die impliziten Erwartungen bzw. das Vertrauen in die Automation und die Neigung zur Bildung eines Übervertrauens minimiert werden.

## *2.12 Zusammenfassung der Theorie*

Unterschiedlichste Studien konnten zeigen, dass Fahrer umso weniger Situationsbewusstsein haben, desto weniger sie direkt in den Regelkreis der Fahrzeugführung eingebunden sind. Ebenso neigen Fahrer dazu, sich selbst, durch falsche implizite Annahmen über die Leistungsfähigkeit einer Automation (Übervertrauen), unzulässig aus dem Regelkreis zu nehmen, bzw. ein mangelndes Situationsbewusstsein gegenüber dem aktuellen bzw. zukünftigen Verhalten der Fahrzeugautomation zu haben. Dieser Aspekt des Situationsbewusstsein über das zu erwartende Verhalten der Automation wird auch als Automationsbewusstsein bezeichnet. Dieser Mangel an Situationsbewusstsein durch mangelndes

Automationsbewusstsein, führt zu kritischen Situationen, wenn die Fahrzeugautomation an Systemgrenzen kommt und der Fahrer dies weder grundsätzlich erwartet noch kurzfristig ein Situationsbewusstsein aufbauen kann. Möglichkeiten solche Systemgrenzen grundsätzlich erwartbarer zu machen können in der Anpassung des Vertrauens des Fahrers hin zu einer realistischen Vorstellung über die Leistungsfähigkeit der Automation liegen.

Soll sich der Fahrer aber temporär von der Fahraufgabe abwenden können, also auch explizit das Vertrauen haben dürfen, sich für einige Zeit abwenden zu können, muss die Fähigkeit zur kurzfristigen Wiedererlangung des Situationsbewusstseins, initiiert durch eine Automationsrückmeldung, gegeben sein. Dies ist z.B. im Automationsgrad hochautomatisiertes Fahren (z.B. BASt Definition) der Fall, aber grundsätzlich auch überall da, wo der Fahrer, z.B. durch Übervertrauen, eine mangelnde oder falsche Erwartung von Systemgrenzen hat.

In einer Übersicht diverser Studien konnte im Mittel eine Zeit von etwa 8 Sekunden identifiziert werden, die es braucht, bei einem Fahrer ein Situationsbewusstsein aufzubauen, das ausreichend ist, die Kontrolle über ein hochautomatisiertes Fahrzeug zurück zu übernehmen. Für eine sichere Kontrollübernahme muss also idealerweise mindestens 8 Sekunden bevor sich die Situation zu weit zugespitzt hat, durch die Automation eine Kontrollübernahme initiiert werden.

Aufgrund des hochdynamischen Charakters vieler Verkehrssituationen, kann eine solche 8-sekündige Zeitperiode mit ausgeprägter Vorhersageunsicherheit einhergehen. Dies bedeutet, dass, um tatsächlich alle Situationen zu inkludieren, in denen tatsächlich eine Kontrollübernahme notwendig wird, viele Fehlalarme des Systems in Kauf genommen werden müssen.

In unterschiedlichsten Anwendungskontexten konnten bereits die negativen Effekte einer Häufung von Fehlalarmen, welche als „False Alarm Effect“ oder „Cry Wolf Effect“ bezeichnet werden, bis hin zur vollständigen Missachtung solcher Alarme, beobachtet werden.

Der Cry Wolf Effekt ist ein Lerneffekt, welcher durch die wiederholt erlebte Kombination aus Alarmereignis und dem Ausbleiben des eigentlich erwarteten Grundes des Alarms entsteht.

Eine geeignete Grundlage für die Erklärung der beim Cry Wolf Effekt stattfindenden Lernprozesse stellt die Theorie der Evidenzevaluation dar. Diese wird zu den kognitiven Lerntheorien des kausalen Lernens gezählt.

Hierbei wird die An – oder Abwesenheit bestimmter Aspekte, zur bestärkenden oder schwächenden Evidenz einer vorher angenommenen Zusammenhangshypothese.

Im Kontext des Cry Wolf Effektes wird ein regelmäßiges Ausbleiben eines Automationsfehlers nach einer Kontrollübernahmeaufforderung, bzw. einer Warnung,

als negative Evidenz der Annahme, dass die Warnung verlässlich vor Systemgrenzen warnt, angesehen.

Zentrale Annahme dieser Arbeit ist, dass die Art der Gestaltung eines „Alarms“ bzw. der Automationsrückmeldung über Ihren Zustand, die Art der Zusammenhangshypothese bereits so beeinflusst, dass diese nicht oder nur schwer widerlegbar wird, und dadurch eine gewisse Resistenz gegenüber einem Cry Wolf Effekt erzeugt werden kann.

Die Rückmeldung von Automationsunsicherheit scheint dafür geeignet zu sei, da Unsicherheit „ergebnisoffen“ interpretiert werden kann und sowohl das Ausblieben als auch das Eintreten eines Automationsfehlers nach einer Warnung in das Konzept der Unsicherheit passt.

Das Konzept der Unsicherheit wurde in verschiedenen Formen bereits in Kontext der Warnung vor Automationsfehlern genutzt. Explizit im Kontext der Vermeidung von Cry Wolf Effekten sind sogenannte Likelihoodalarm-Strategien zu nennen. In denen eine Alarmereignis direkt in wahrscheinliche und weniger wahrscheinliche Stufen unterteilt wird (z.B. Safe, unlikely und likely), und diese Stufen zurückgemeldet werden.

Studien konnten zeigen, dass diese Strategien zu einer Verringerung des Cry Wolf Effektes bei den Alarmen führt, die als „Likely“ eingestuft wurden. Was aus den dazu zur Verfügung stehenden Studien im Regelfall nicht hervorgeht, ist eine Aussage zur Verschlechterung der Performanz bei „Unlikely“ Ereignissen. Denn auch hier ist es ja nicht ausgeschlossen, dass es zu Automationsfehlern kommt. Die Performanz der „Unlikely-Stufe“ würde, so gesehen, zugunsten der Performanz der „Likely-Stufe“ geopfert werden.

Unsicherheit wird auch in anderen Ansätzen zurückgemeldet. Jedoch wurden in diesen Untersuchungen keinerlei Auswirkungen auf den Cry Wolf Effekt betrachtet. So wird z.B. direkt den Grad der Unsicherheit des Systems bezüglich bestimmter „Diagnosen“ (z.B. in der automatisierten Bildinterpretation oder in der Aussage über bestimmte Zustände eines Flugzeuges zurückgemeldet.

Auch im Fahrzeug wurden Strategien untersucht, eine Fahrzeugunsicherheit kontinuierlich zurückzumelden. Es konnten hier positive Effekte hinsichtlich der Anpassung des Verhaltens, bzw. der Erwartung von Automationsfehlern gefunden werden. Zu nennen wäre hier zum Beispiel eine bessere Akzeptanz von Automationsfehlern, da vorher in Form der Rückmeldung der Fahrzeugunsicherheit eine Art Begründung für das Automationsverhalten geliefert wurde.

### 3 Leithypothesen

Im Theorieteil werden zwei unterschiedliche Ansatzpunkte zur Erhöhung der Erwartung von Systemgrenzen und Automationsfehlern erörtert, die möglicherweise durch eine Unsicherheitsrückmeldung realisierbar sind.

- Die situative Erhöhung der Erwartung von Systemgrenzen und daraus möglicherweise folgenden Automationsfehlern durch gezielt platzierte Rückmeldungen der Automation, z.B. in Form von Warnungen / Alarmen oder eben Unsicherheit. Bei Maßnahmen zur situativen Erhöhung der Erwartung von Automationsfehlern besteht grundsätzlich die Möglichkeit eines Cry Wolf Effektes.
- Die Beeinflussung der generellen Erwartung von Systemgrenzen und Automationsfehlern implizit durch die Anpassung des Vertrauens in die Automation.

Zur Untersuchung möglicher Effekte auf diese beiden Bereiche, werden Leithypothesen verfasst.

#### 3.1 *Auswirkungen auf die situative Erwartung von Automationsfehlern:*

- **Leithypothese 1 und 1. notwendiges Kriterium:** Die Rückmeldung von Automationsunsicherheit ist grundsätzlich effektiv bezüglich der Erhöhung der Erwartung eines Automationsfehlers. Daher werden Automationsfehler, vor denen Unsicherheit zurückgemeldet wird, besser kontrolliert als Automationsfehler, vor denen keine Rückmeldung stattfindet.
- **Leithypothese 2 und 2. notwendiges Kriterium:** In Abhängigkeit der Häufigkeit des erlebten Ausbleibens eines Automationsfehlers nach einer Rückmeldung lässt sich ein Cry Wolf Effekt erzeugen.

Wenn vor einem tatsächlichen Automationsfehler im vorherigen Verlauf der Fahrt eine Reihe von Rückmeldungen ohne nachfolgenden Automationsfehler erlebt wurde, entsteht ein Cry Wolf Effekt. Dieser äußert sich u.a. in einer schlechteren Kontrollierbarkeit des Automationsfehlers in Vergleich zu keinem Cry Wolf Effekt. Dabei ist die Stärke des Cry Wolf Effektes eine Funktion der Anzahl der in einer Reihe erlebten Rückmeldungen ohne nachfolgenden Automationsfehler.

- **Leithypothese 3:** Über die Art der Rückmeldungsgestaltung kann auf die Art der Zusammenhangshypothese zwischen einer Rückmeldung und dem Eintreffen bzw. Ausbleiben eines (Automationsfehler)Ereignisses Einfluss genommen werden. Dadurch kann die Neigung zur Entstehung eines Cry Wolf Effektes beeinflusst werden:

Wird die Theorie der Evidenzevaluation zugrunde gelegt (siehe theoretischer Hintergrund), ist die Zusammenhangshypothese eines konventionell gestalteten Alarms, „starr“ bzw. unflexibel in dem Sinne, dass sie das Nicht-Eintreffen eines negativen Ereignisses nicht vorsieht. Das Ausbleiben der erwarteten Konsequenz stellt negative Evidenz für die Gültigkeit der Zusammenhangshypothese dar. Ein Cry Wolf Effekt kann demnach durch genügend häufiges Erleben von „falschen Alarmen“ bzw. negativen Evidenzen konditioniert werden, da der Zusammenhang zwischen ursächlichem Ereignis und Alarm nicht mehr als zuverlässig bzw. stabil genug angesehen wird. Eine Gestaltung der Rückmeldung als Automationsunsicherheit erzeugt keine verbindliche Zusammenhangshypothese zwischen einer Rückmeldung und möglichen Konsequenzen. Das Ausbleiben eines Automationsfehlers nach einer Rückmeldung kann folglich nicht als falscher Alarm bezeichnet werden, da dies explizit Teil der Zusammenhangshypothese ist. Eine Konditionierung eines Cry Wolf Effektes ist in diesem Fall entweder gar nicht möglich, oder die Cry Wolf Konditionierung ist deutlich schwerer möglich, da im Vergleich zu konventionell gestalteten Alarmen deutlich mehr Rückmeldungen ohne Automationsfehler erlebt werden müssen, um eine Art von negativer Zusammenhangshypothese zu erzeugen.

- **Leithypothese 4:** Das Erleben positiver Evidenz stabilisiert die Zusammenhangshypothese und wirkt sich in Abhängigkeit der Alarmgestaltung unterschiedlich aus.

Wird bei einem konventionellen Alarm nach einer Rückmeldung ein Automationsfehler erlebt, so wirkt sich dies hemmend auf die Entstehung zukünftiger Cry Wolf Effekte aus.

Das heißt die Alarmeffektivität und damit verbunden die Kontrollierbarkeit ist beim ersten Automationsfehler, der erlebt wird, niedriger als beim zweiten erlebten Automationsfehler, wenn vor beiden Automationsfehlern die gleiche Anzahl an falschen Alarmen erlebt wurde.

Bei einer Unsicherheitsrückmeldung stellen das Eintreten als auch das Ausbleiben eines Automationsfehlers nach der Rückmeldung gleichartige

Optionen dar. Daher stellt das Ausbleiben keine negative Evidenz – das Eintreten aber auch keine positive Evidenz dar. Folglich hat das Erleben von echten Automationsfehlern bei Unsicherheitsrückmeldung einen deutlich weniger hemmenden Einfluss auf die Entstehung eines Cry Wolf Effektes als bei einem normalen Alarm. Es gibt einen Interaktionseffekt aus Rückmeldungsart und Fehlererfahrung.

### **3.2 Auswirkungen auf die generelle Erwartung von Automationsfehlern**

- **Leithypothese 5:** Vermeidung der Entstehung von Overtrust

Das Erleben von echten Automationsfehlern kann das Vertrauen in die Automation anpassen, da diese Erfahrung die Einschätzung der Leistungsfähigkeit der Automation beeinflusst. Wird ein Fehler erlebt, so kann dies über die Anpassung des Vertrauens die generelle Erwartung eines Fehlers fördern. Die alleinige Rückmeldung eines Automationsstatus, der auf Automationsfehler hinweist, wirkt der Entstehung von Overtrust entgegen, auch wenn die Automation im erlebten Zeitraum gar keine tatsächlichen Fehler macht. Nach einer Fahrt nur mit Unsicherheitsrückmeldung aber ohne Automationsfehler ist das Vertrauen in die Automation niedriger als nach einer Fahrt ohne Automationsfehler aber ohne Rückmeldungen.

- **Leithypothese 6:** Verhinderung von Misstrust durch Erhöhung der Vorhersehbarkeit von Automationsfehlern

Das Erleben von Automationsfehlern kann zur Entwicklung eines Misstrauens in die Automation führen. Dies gilt insbesondere dann, wenn Fehler nicht systematisch vorhergesagt werden können und jederzeit mit einem Fehler gerechnet werden muss. Insbesondere das Erleben von gar nicht erwarteten Fehlern, auch gerade aus einem Overtrust heraus, kann zu einem Kippen des Vertrauens in Richtung des Misstrust führen, da aufgrund der mangelnden Erwartung des Fehlers dieser schlechter kontrolliert werden kann.

Die situative Erhöhung der Erwartung eines Automationsfehlers durch Rückmeldungen, kann situativ die Einschätzung der Leistungsfähigkeit einer Automation anpassen, was der Entstehung von Misstrust der Automation im Allgemeinen entgegenwirkt. Diese Wirkebene wird stark durch die Reliance in



das Rückmeldesystem / Alarmsystem beeinflusst und ist daher stark durch das Vorhandensein und die Stärke eines Cry Wolf Effektes beeinflusst. Ist die Reliance in das Rückmeldesystem gering, so sinkt auch der Effekt auf die situative Erhöhung der Fehlererwartung. Nach einer Fahrt mit Automationsfehler ist das Vertrauen in die Automation daher niedriger, wenn zum Zeitpunkt des Automationsfehlers ein Cry Wolf Effekt vorlag. Unsicherheitsrückmeldung führt in diesem Fall zu einem besseren Vertrauen als ein normaler Alarm.

- **Leithypothese 7 und 3. notwendiges Kriterium.** Das Vorhandensein eines Systems das Automationsunsicherheit vor Automationsfehlern zurückmeldet, darf, wenn vor dem Fehler eine Rückmeldung ausbleibt, nicht zu einer geringeren Erwartung eines Automationsfehlers und somit zu einer schlechteren Kontrollierbarkeit führen als in einer Bedingung ganz ohne Rückmeldesystem.

Hier soll im Kontext der Unsicherheitsuntersuchung betrachtet werden in wie weit der Grad der Überzeugung, dass ein Alarmsystem nur dann anschlägt wenn auch ein tatsächlicher Grund vorliegt (Compliance), den Grad der Überzeugung, dass ein tatsächlich vorhandener Fehler auch durch das Rückmeldesystem abgedeckt wird (Reliance) beeinflusst (z.B. Meyer, 2001 & 2004). Wenn eine Abhängigkeit zwischen beiden bestehen sollte – würde dies bedeuten, dass Alarmsystem mit einer hohen Compliance, also auch einer niedrigen Cry Wolf Neigung – schlechter sind bezüglich der Erwartung von Fehlern, die nicht durch das System vorhergesagt werden. Im Umkehrschluss würde dies bedeuten, dass ein Alarmsystem mit einem hohen Cry Wolf Effekt – zu einer höheren Erwartung von Automationsfehlern, vor denen kein Alarm bzw. Rückmeldung erfolgt, führt. Eine Abhängigkeit von Compliance und Reliance wurde in einem anderen Kontext bereits beobachtet (Wickens & McCarly, 2006 & 2007)

## **4 Übersicht der Studien und Zuordnung zu Leithypothesen**

Die insgesamt 7 aufgeführten Leithypothesen werden innerhalb der 3 für diese Arbeit durchgeführten Studien überprüft. Dafür werden diese entweder direkt als Hypothesen übernommen, oder, durch die Formulierung von Unterhypothesen, weiter operationalisiert. Zur Überprüfung der Leithypothesen wurde ein übergeordnetes Versuchsparadigma entwickelt, das über alle 3 Teilstudien

eingesetzt wurde. Dieses Vorgehen wird im Teil „Basis-Untersuchungsparadigma“ ausführlich erläutert, die wesentlichsten Aspekte sollen hier aber, zum besseren Verständnis des Studienüberblicks, kurz umrissen werden.

Alle Studien wurden in simulierten Fahrten in einem einfachen fixed base Simulator durchgeführt. In allen Studien fuhren Versuchspersonen mit einer Automationsunterstützung, einer kombinierten Längs- und Querführung und sollten dabei eine Nebenaufgabe bearbeiten. Die Nebenaufgabe sollte Ablenkung erzeugen, um die Wirkung von Automationsalarman auf die Verteilung der Aufmerksamkeit zwischen Fahr und Nebenaufgabe erfassen zu können. Während der Fahrt gab es immer wieder Situationen, in denen das Fahrzeug einem Fremdfahrzeug begegnete. In diesen Situationen reagierte die Automation entweder vollkommen korrekt und es wurde auch kein Automationsalarm gegeben, oder die Automation reagierte korrekt und es wurde vorher ein, je Studie/Bedingung variierender Automationsalarm gegeben, es resultierte also ein falscher Alarm. Ebenso gab es Situationen, in denen die Automation nicht korrekt reagierte und, zur Vermeidung eines Unfalls, ein Fahrereingriff erforderlich war. Auch hier gab es entweder eine Automationsrückmeldung oder keine. Innerhalb der einzelnen Teilstudien wurde das Untersuchungsparadigma den jeweiligen Studienzielen entsprechend variiert. Im Folgenden wird eine kurze Übersicht über diese Studien gegeben.

## *4.1 Übersicht Studie 1*

Studie 1 hat einen teils explorativen Charakter und soll in erster Linie dazu dienen, erste Eindrücke zur Unsicherheitsrückmeldung gewinnen zu können.

Im Wesentlichen werden in Studie 1 folgende Bedingungen hinsichtlich ihrer Haupt- und Interaktionseffekte überprüft:

- Unsicherheitsrückmeldung vs. keine Rückmeldung bezüglich der Kontrollierbarkeit eines Automationsfehlers, bei dem der Fahrer nur durch einen Bremseneingriff einen Auffahrunfall verhindern kann. Diese Betrachtung der Kontrollierbarkeit diene der Überprüfung von **Leithypothese 1**: Eine Rückmeldung einer Automationsunsicherheit ist grundsätzlich effektiv bezüglich der Erhöhung der Erwartung eines Automationsfehlers.
- Des Weiteren wurde die Anzahl der erlebten Unsicherheitsrückmeldungen ohne nachfolgenden Automationsfehler (falsche Alarmer) variiert, und dadurch ein Cry Wolf Effekt Potenzial aufgebaut. Die Ausprägung eines möglichen Cry Wolf Effektes wurde dann anhand der Kontrollierbarkeit eines schließlich doch folgenden Automationsfehlers überprüft. Dafür wurde entweder früh in einer

Fahrt, ohne vorher falsche Alarme erlebt zu haben, eine Automationsfehler erzeugt, oder erst spät in einer Fahrt, nachdem schon eine Reihe falscher Alarme erlebt wurden. Dies diente zur Überprüfung der **Leithypothese 2**: Durch Variation erlebter Automationsrückmeldung ohne Automationsfehler lässt sich ein Cry Wolf Effekt erzeugen.

- Mittels eines Vergleichs einer Fahrt in der zwar Unsicherheitsrückmeldung aber keine tatsächliche Automationsfehler erlebt wurden, mit einer Fahrt in der Unsicherheitsrückmeldung und anschließende Automationsfehler erlebt wurden, sollte **Leithypothese 5** überprüft werden: durch die Betrachtung, ob allein die Rückmeldung einer Automationsunsicherheit, ohne tatsächliches Erleben eines Automationsfehlers, das Vertrauen in die Automation so anpassen kann, dass auch Fehler erwartet werden.
- Ein weiterer Vergleich wurde angestellt, zwischen dem Erleben von Automationsfehlern mit vorheriger Unsicherheitsrückmeldung vs. dem Erleben von Automationsfehlern ohne vorherige Unsicherheitsrückmeldung. Durch diesen Vergleich sollte **Leithypothese 6** überprüft werden, dass eine Rückmeldung vor einem Automationsfehler sich dämpfend auf die Entwicklung von Misstrauen in die Automation auswirkt, verglichen mit keiner Rückmeldung. Hierzu wurden, durch weitere Variation des Zeitpunkts des ersten Erlebens eines Automationsfehlers, früh (kein Overtrust entsteht) vs. spät in einer Fahrt (Tendenz zu Overtrust), mögliche moderierende Einflüsse des vor dem Automationsfehler aufgebauten Übervertrauens betrachtet. Je weniger ein Automationsfehler erwartet wird (spät / Overtrust), desto größer ist das Potenzial durch einen Automationsfehler Misstrauen in die Automation zu erzeugen.

Die Betrachtung der subjektiv empfundenen Kooperation zwischen Fahrer und Fahrzeug ermöglichte die Überprüfung einer weiteren Facette der Leithypothese 6, in wie weit eine Unsicherheitsrückmeldung zu einer besseren Akzeptanz einer nicht fehlerfreien Automation beiträgt.

## 4.2 Übersicht Studie 2

Studie 2 stellt eine Kombination unterschiedlicher Teilexperimente dar, in denen jeweils der Fokus auf den Vergleich bestimmter Bedingungen gelegt wird. In Studie 2 werden folgende Bedingungen hinsichtlich Ihrer Haupt- und Interaktionseffekte überprüft:

- Vergleich der Unsicherheitsrückmeldung mit einem konventionell gestalteten Alarm. Der konventionelle Alarm ist dabei in der Art gestaltet, wie man ihn

häufig im Kontext von Kontrollübernahmeaufforderungen in Fahrzeugen findet. Ähnlich wie in einigen Bedingungen in Studie 1, sollte durch das wiederholte Erleben falscher Alarme, eine Neigung zu einem Cry Wolf Effekt evoziert werden.

Der Vergleich diene in erster Linie wesentliche Facetten der **Leithypothese 3** zu überprüfen, dass durch die Art der Alarmgestaltung die Zusammenhangshypothesen zwischen Alarm und erwarteter Folge bzw. Ursache so beeinflusst werden kann, dass ein Cry Wolf Effekt vermieden / verringert werden kann. Die Effekte der Alarmgestaltung und deren Wechselwirkung mit anderen Faktoren, wurde anhand mehrerer abhängigen Variablen erfasst.

- Die Kontrollierbarkeit in einer „echten“ Automationsfehlersituation nach einem Alarmereignis.
  - Die Auswirkungen auf das subjektiv empfundene situative Vertrauen in die Automation bzw. die situative Erwartung, ob die Automation einen Fehler macht oder nicht.
  - Die situative Veränderung der Verteilung der Aufmerksamkeit zwischen Fahraufgabe und einer Nebenaufgabe in der Zeit kurz nach einer Automationsrückmeldung. Dies wurde auch als weiteres Indiz für das situative Vertrauen herangezogen.
- Der Vergleich der Unsicherheitsrückmeldung mit einem konventionellen Alarm wurde auch genutzt, um **Leithypothese 6** zu überprüfen: Kann ein Cry Wolf Effekt eine negative Auswirkung auf das Vertrauen in die Automation als ganze haben, also Misstrauen erzeugen, bzw. wirkt sich eine Unsicherheitsrückmeldung in irgendeiner Weise schützend auf die Vermeidung von Misstrauen in die Automation aus? Dies wurde durch die Erhebung des Gesamtvertrauens in die Automation nach den Fahrten erfasst.
- Vergleich unterschiedlich hoher Cry Wolf Potenziale erzeugt durch Variation der Menge, der in unterschiedlichen Fahrten erlebten „falschen Alarme“. Hierdurch sollte im Wesentlichen bezüglich **Leithypothese 3** überprüft werden, in wie weit die Häufigkeit bzw. Menge der erlebten falschen Alarme einen moderierenden Einfluss auf die Neigung einen Cry Wolf Effekt bei den verschiedenen Rückmeldestrategien haben kann. Der Vergleich diene aber auch der weiteren Überprüfung des notwendigen Kriteriums der **Leithypothese 2**, dass durch eine Variation der Menge der erlebten falschen Alarme ein Cry Wolf Effekt überhaupt provoziert werden kann.

### 4.3 Übersicht Studie 3:

Auch Studie 3 stellt eine Kombination unterschiedlicher Teilexperimente dar, in denen jeweils der Fokus auf den Vergleich bestimmter Bedingungen gelegt wird. In Studie 3 werden folgende Bedingungen auf deren Haupt- und Interaktionseffekte bezüglich unterschiedlicher abhängiger Variablen überprüft.

- Der Vergleich der Unsicherheitsrückmeldung mit einem konventionell gestalteten Alarm aus Studie 2 wurde wiederholt, um eine in Studie 2 identifizierte Sekundärvarianz kontrollieren zu können. Hier zeigte sich, dass die Gleichartigkeit der in Studie 2 erlebten Fahrsituationen dazu führte, dass Versuchspersonen die Situation an sich als nicht mehr gefährlich beurteilten und somit nicht mehr klar identifizierbar war, ob die Varianz in den abhängigen Variablen nun durch einen Cry Wolf Effekt, bzw. die Alarmgestaltung, oder aufgrund der Bewertung der Situation an sich zustande kam. Wie in Studie 2, wurde durch diesen Vergleich wieder **Leithypothese 3** (durch die Art der Alarmgestaltung kann die Zusammenhangshypothesen zwischen Alarm und erwarteter Folge bzw. Ursache so beeinflusst werden, dass ein Cry Wolf Effekt vermieden bzw. abgeschwächt werden kann) überprüft.
- Durch den weiteren Vergleich der Unsicherheitsrückmeldung mit einem idealen Alarm, der nur dann eine Rückmeldung gibt, wenn tatsächlich ein Fahrereingriff nötig ist, sollte die absolute Größe eines möglichen Cry Wolf Effektes einer Unsicherheitsrückmeldung bestimmt werden. Was eine weitere Facette der **Leithypothese 3** darstellt.
- Die Unsicherheitsrückmeldung, sowie auch die anderen Rückmeldungen wurden mit einer Mischstrategie im Sinne der sogenannten Likelihoodalarmstrategie verglichen, in der in Situationen ohne Automationsfehler Unsicherheit und in Situationen mit Automationsfehler ein konventioneller Alarm zurückgemeldet wurde. Dieser Vergleich ist am ehesten **Leithypothese 3** zuzuordnen, diente aber in erster Linie dazu, eine Aussage treffen zu können, ob ein Alarm/Unsicherheitsrückmeldung in mehr als zwei Stufen gestaltet werden sollte.
- Die oben genannten unterschiedlichen Rückmeldungsarten wurden in einem Vergleich zwischen einer ersten Fahrt, bis zum ersten Automationsfehler und einer Zweiten Fahrt, ab dem ersten erlebten Automationsfehler, betrachtet. Dies sollte die **Leithypothese 4** überprüfen: Das Erleben positiver Evidenz stabilisiert die Zusammenhangshypothese und wirkt sich in Abhängigkeit der Alarmgestaltung unterschiedlich aus. Leithypothese 4 bildet dabei einen Unteraspekt der Leithypothese 3. So wird hier z.B. vermutet, dass je höher der Cry Wolf Effekt beim ersten Automationsfehler ist, desto stärker wirkt sich das

Erleben dieser „positiven Evidenz“ hemmend auf die weitere Entwicklung eines Cry Wolf Effektes aus.

- Des Weiteren wurde ein Vergleich angestellt, wie sich die verschiedenen Rückmeldungsarten in einer Situation mit Automationsfehler, aber ohne vorherige Warnung – im Vergleich zu vorher erlebten Situationen eines Automationsfehlers mit vorheriger Automationsrückmeldung, auf die Kontrollierbarkeit des Automationsfehlers auswirken. Dies sollte zur Überprüfung der **Leithypothese 7** dienen: Das Vorhandensein eines Systems das Automationsunsicherheit vor Automationsfehlern zurückmeldet, darf, wenn vor dem Fehler eine Rückmeldung ausbleibt, nicht zu einer geringeren Erwartung eines Automationsfehlers und somit zu einer schlechteren Kontrollierbarkeit führen, als in einer Bedingung ganz ohne Rückmeldesystem.

Das Untersuchungsparadigma und die über alle Teilstudien verwendeten Versuchsmaterialien werden im folgenden Teil vorgestellt.

## **5 Basis-Untersuchungsparadigma**

In diesem Kapitel wird das grundlegende Untersuchungsparadigma dargestellt welches für alle Teilstudien als Ausgangspunkt genutzt und dort entsprechend variiert wird. Die für die einzelnen Studien relevanten Auslegungen des Grundparadigmas werden in den jeweiligen Abschnitten der Teilstudien erläutert. Teilziel der Arbeit ist die Erarbeitung eines Untersuchungsparadigmas, das es ermöglichen soll, die moderierenden Einflüsse der Rückmeldungsgestaltung auf das Erleben und Verhalten der Fahrer, insbesondere das Vertrauen in die Automation und die Anfälligkeit für einen Cry Wolf Effekt, beobachtbar zu machen.

Das auszuwählende Paradigma muss geeignet sein in der relativ geringen Zeit, in der die Fahrer sich mit dem System auseinandersetzen, sowohl ein Potenzial für einen Cry Wolf Effekt zu erzeugen als auch das Vertrauen in die Automation zu beeinflussen.

### ***5.1 Methodik zur Untersuchung der Cry Wolf Anfälligkeit***

Ein Cry Wolf Effekt kann entstehen, wenn ein Diagnose- bzw. Alarmsystem zu häufig falsche Alarme produziert. Dies bedeutet, dass ein Ereignis, welches das System scheinbar identifiziert hat, in Wirklichkeit gar nicht vorliegt. Zur Untersuchung der Cry Wolf Anfälligkeit unterschiedlicher Rückmeldestrategien im Kontext des automatisierten Fahrens ist es daher zunächst nötig, Situationen zu schaffen in denen solche falschen Alarme einer Fahrzeugautomation durch den Fahrer erlebt

werden können. Ziel ist es ein Fahrscenario zu erstellen, das zum einen durchaus gängige und plausibel erscheinende Automationsfehler beinhaltet, aber auch Situationen umfasst, die Fehlalarme nicht vollkommen willkürlich, sondern in bestimmtem Maße situationsbezogen erscheinen lässt.

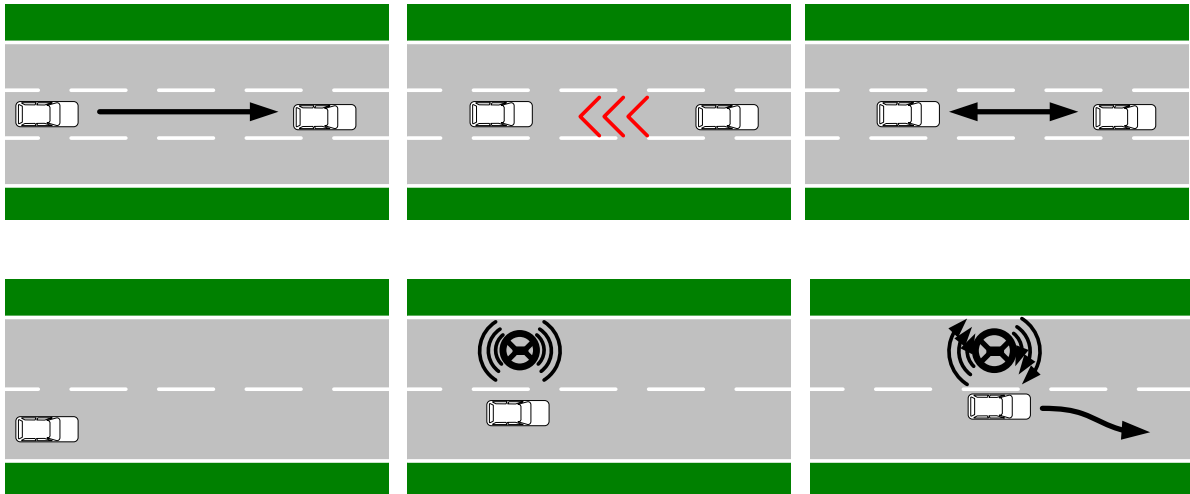
Ausgangspunkt des Grundparadigmas ist die Fahrt in einem simulierten Fahrzeug mit einer Kombination aus Längs- und Querführungsunterstützung über eine gerade Autobahn mit zwei Fahrstreifen in eine Richtung. Die Automation soll so gestaltet sein, dass sie die im Theorieteil beschriebenen Effekte des Übervertrauens und daraus folgender Verhaltensanpassung, wie z.B. vermehrte Abwendung von der Fahraufgabe, provozieren kann. Zugleich sollte sie den Fahrer aber nicht vollkommen aus der Verantwortung über die Fahraufgabe entlassen. Nach den Definitionen der Automationsgrade nach Gasser et al. (2012), handelte es sich um eine Mischung aus assistiertem und teilautomatisiertem Fahren.

Die Art der eingesetzten Automation soll sich von der Bedienung und der Auslegung nicht übermäßig von bereits in Serie erhältlichen Systemen unterscheiden. Daher wird eine Kombination aus einem adaptiven Abstandsregeltempomaten (Adaptive Cruise Control – ACC) und einer leichten Spurhalteunterstützung umgesetzt (Abbildung 9).

Die Versuchspersonen sollen, unter Einsatz der Fahrzeugautomation, auf der linken Spur fahren und dabei möglichst die in der Automation voreingestellte Geschwindigkeit von 100 km/h beibehalten. Während der Fahrt passieren in unregelmäßigen Abständen Situationen, in denen ein 50 km/h schnelles Fremdfahrzeug auftaucht.

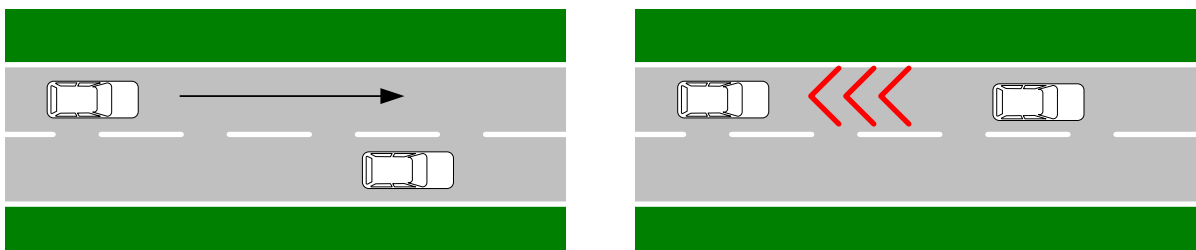
Die Fahrten finden im Nebel bei 90 Meter Sichtweite statt. Der Nebel wird benötigt, da aus softwaretechnischen Gründen die Fremdfahrzeuge, je nach Fahrsituation, in einem Abstand von 90 Metern zum Ego-Fahrzeug in das Fahrscenario hineingesetzt werden und ohne Nebel für den Fahrer somit einfach aus dem Nichts auftauchen würden.

Um die voreingestellten 100 km/h im Nebel für die Versuchspersonen nicht als zu schnell und daher unnötig gefährlich erscheinen zu lassen, wird die Sichtweite des Nebels so gewählt, dass innerhalb der Nebelsichtweite ein Abbremsen bis zum Stillstand möglich ist. Bei einer Reaktionszeit von 1 Sekunde, einer Schwellzeit von 0,2 Sekunden und einer Verzögerung von  $8 \text{ m/s}^2$  (bei trockener, griffiger Fahrbahn), kann von einem Anhalteweg von etwa 77 Metern ausgegangen werden.



**Abbildung 9: Schematische Darstellung der Automationsfunktion. Oben:** Längsführungsfunktion, bei Annäherung an ein langsamer als die eingestellte Geschwindigkeit des eigenen Fahrzeuges fahrendes Fahrzeug, erfolgte ein automatisches Abbremsen bis die Relativgeschwindigkeit gleich null war, danach wurde dem Vorderfahrzeug mit einem bestimmten Abstand gefolgt. **Unten:** Quersührungsfunktion, bei Annäherung an eine seitliche Spurmarkierung erfolgte zunächst eine Vibration auf dem Lenkrad, die bei weiterer Annäherung um eine Gegenkraft in Richtung der Fahrbahnmitte ergänzt wurde.

Die meisten Situationen mit einem Fremdfahrzeug sind Situationen mit einem normalen, der Situation angemessenen Automationsverhalten. In diesen Situationen wird keine Rückmeldung der Automation geäußert. Diese Situationen werden als „klare“ Situationen bezeichnet. In klaren Situationen fahren die Fremdfahrzeuge entweder deutlich auf dem Nachbarfahrstreifen oder deutlich auf dem eigenen Fahrstreifen. In diesen Situationen hat die Automation kein Problem, die anderen Fahrzeuge zu detektieren und reagiert der Situation immer angemessen. Das heißt, dass sie bei einem Fremdfahrzeug auf dem eigenen Fahrstreifen abbremst und in die Folgefahrt geht (Abbildung 10 – rechts) und bei einem Fremdfahrzeug auf dem Nachbarfahrstreifen an diesem vorbeifährt (Abbildung 10 - links).



**Abbildung 10: exemplarische Darstellung „klarer“ Fahrsituationen.**

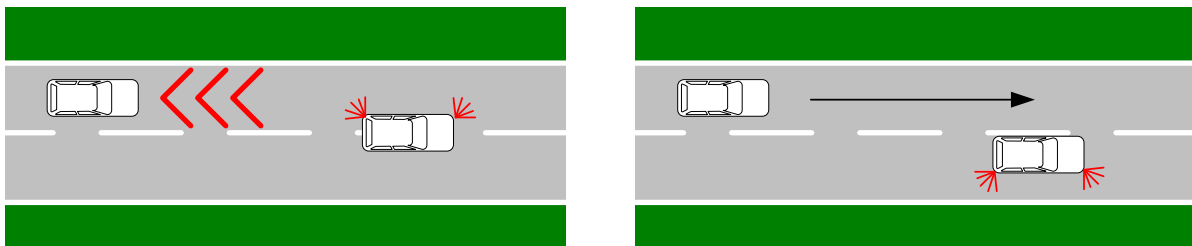
Zur Provokation eines Cry Wolf Effektes werden in diese Fahrt, je nach Untersuchungsschwerpunkt, unterschiedlich viele Situationen eingestreut, in denen



die Automation eine Rückmeldung äußert – es aber danach nicht zu einem Automationsfehler, sondern zu einer der Situation angemessenen Situation kommt. Situationen, in denen eine Rückmeldung geäußert wird, sollten sich allerdings von den regulären unkritischen Situationen insofern unterscheiden – dass sie auch für den Fahrer keine eindeutige Interpretation der Situation zulassen. Diese Fahrsituationen werden daher „unklare“ Fahrsituationen genannt. In diesen unklaren Situationen fahren die Fremdfahrzeuge entweder auf dem rechten Fahrstreifen so nach links versetzt, dass die Spurmarkierung zum Fahrstreifen des Ego-Fahrzeuges fast überfahren wird, oder auf dem linken Fahrstreifen, jedoch deutlich nach rechts versetzt.

In den unklaren Situationen mit angemessenem Automationsverhalten kann die Fahrzeugautomation die anderen Fahrzeuge zwar nicht eindeutig detektieren bzw. einem Fahrstreifen zuordnen, sie reagiert aber dennoch angemessen.

Angemessenes Reagieren ist entweder das Abbremsen und hinter dem versetzt auf dem eigenen Fahrstreifen fahrenden Fremdfahrzeug in die Fahrfahrt wechseln (Abbildung 11 –links), oder das Vorbeifahren an dem nach links versetzt auf dem rechten Fahrstreifen fahrenden Fremdfahrzeug (Abbildung 11 –rechts).

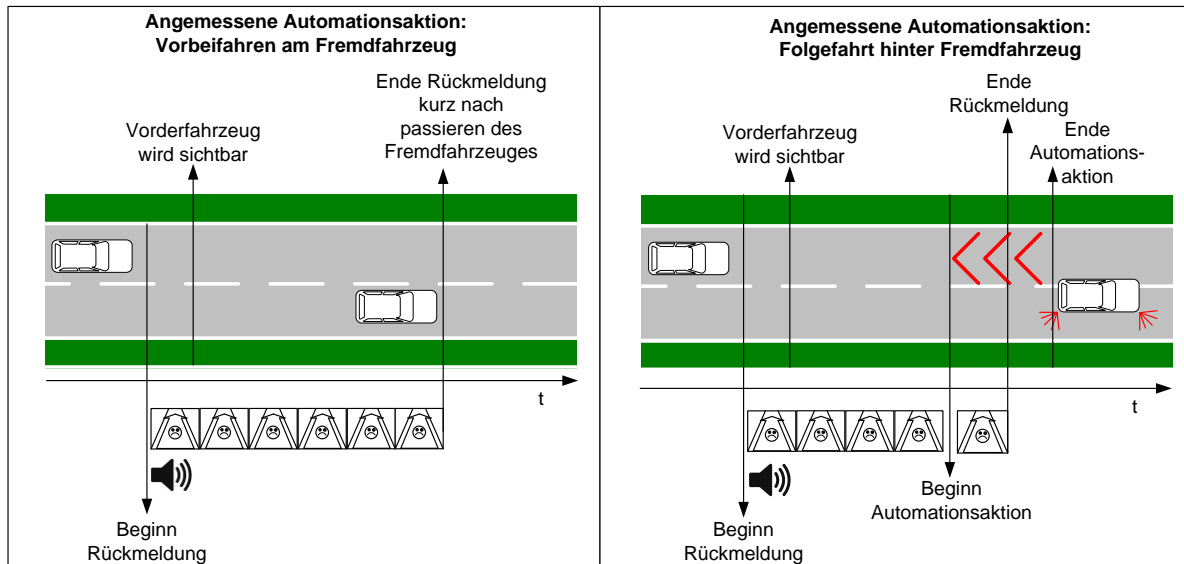


**Abbildung 11: Beispielhafte Darstellung „unklarer“ Fahrsituationen ohne Automationsfehler, links Fahrzeug fährt deutlich versetzt auf dem Fahrstreifen des Ego-Fahrzeuges.**

In den unklaren Fahrsituationen kann eine Rückmeldung der Automation erfolgen, auf die aber kein Automationsfehler folgt. Die Art der Rückmeldung kann dabei je nach Versuchsbedingung variiert werden, das zeitliche Ablaufschema der Rückmeldungen ist aber, unabhängig von der Art der Rückmeldung, immer das Gleiche.

Die Automationsrückmeldung beginnt immer mindestens 2 Sekunden vor dem Sichtbarwerden des Vorderfahrzeuges und wird durch ein akustisches Signal angekündigt. Das akustische Signal dient dazu, die Aufmerksamkeit des Fahrers auf die Automationsrückmeldung zu lenken. So soll sichergestellt werden, dass auch abgelenkte Fahrer die Automationsrückmeldung rechtzeitig wahrnehmen können. Nach dem Sichtbarwerden des Fremdfahrzeuges wird die Rückmeldung für eine Strecke von mindestens 100 Metern, variierbar nach Teilstudie, angezeigt. Die Szenarien in den Teilstudien werden so abgestimmt, dass die Rückmeldung bei

angemessener Automationsaktion, spätestens dann aufhört, wenn die Automation das Fremdfahrzeug überholt hat, oder bei Folgefahrtsituationen, dann wenn die Automation in die Folgefahrt übergang (Abbildung 12).



**Abbildung 12: Ablauf der Automationsrückmeldung in unklaren Fahrsituationen mit angemessener Automationsaktion.**

Alle Fahrsituationen mit einem Fremdfahrzeug, klare als auch unklare Fahrsituationen, haben stets nur eine bestimmte Dauer. Beim Vorbeifahren an Fremdfahrzeugen ist die Situation recht schnell vorbei. Bei Folgefahrten fährt das Fremdfahrzeug etwa 15 Sekunden lang noch auf dem Fahrstreifen des Ego-Fahrzeuges und wechselt danach auf den rechten Fahrstreifen. Die Automation des Ego-Fahrzeuges beschleunigt daraufhin wieder auf 100 km/h und überholt somit das Fremdfahrzeug.

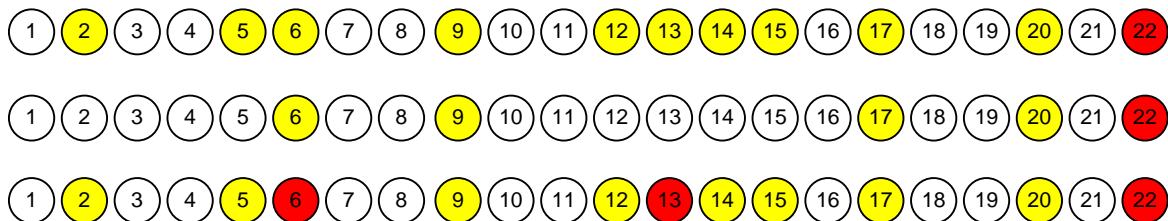
Zwischen den Fahrsituationen mit einem Fremdfahrzeug, gibt es in ihrer Länge nach einem Zufallsprinzip variierte Abschnitte, in denen das Ego-Fahrzeug allein auf der Strecke ist. Situationen mit einem Fremdfahrzeug geschehen also nicht immer in einem immer gleichen Zeitabstand, sondern in einem zwischen 15 bis 60 Sekunden variierenden Abstand. Das Fahrszenario soll durch die Variation der Abstände zwischen den Situationen natürlicher erscheinen. Des Weiteren soll dadurch verhindert werden, dass die durch Automationsrückmeldung erzeugte situative Erwartung eines Ereignisses nicht durch immer gleiche Zeitabstände des Auftretens einer Situation konfundiert wird.

### 5.1.1 Variation der Erfahrung „falscher“ Alarme

Zur systematischen Untersuchung der Cry Wolf Anfälligkeit unterschiedlicher Rückmeldungen kann nun die relative Häufigkeit des Anteils der Situationen mit

Rückmeldung aber ohne Automationsfehler an den Gesamtsituationen variiert werden.

So werden Fahrtbedingungen realisiert, die entweder wenige oder viele Fahrsituationen mit falschen Alarmen auf die jeweilige Gesamtmenge von Fahrsituationen enthalten (eine schematische Darstellung einer beispielhaften Umsetzung zeigt Abbildung 13 obere und mittlere Reihe).

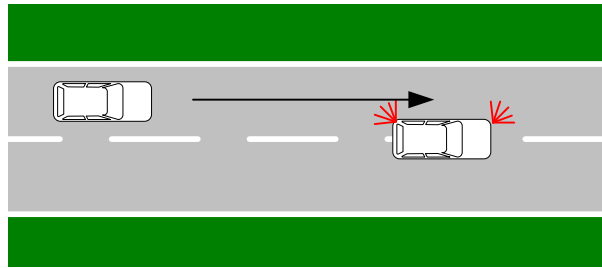


**Abbildung 13: Exemplarische Darstellung der Erzeugung unterschiedlich hohen Cry Wolf Potenzials. Weiße Kreise stellen Fahrsituationen OHNE Rückmeldung und OHNE Automationsfehler dar. Gelbe Kreise stellen Situationen MIT Rückmeldung aber OHNE Automationsfehler dar. Rote Kreise stellen Fahrsituationen MIT Rückmeldung und MIT Automationsfehler dar. Obere Reihe: hohes Cry Wolf Potenzial – Mittlere Reihe: mittleres Cry Wolf Potenzial. – Untere Reihe niedriges Cry Wolf Potenzial durch eingestreutes Erleben positiver Evidenzen.**

### 5.1.2 Variation der Erfahrung tatsächlicher Automationsfehler

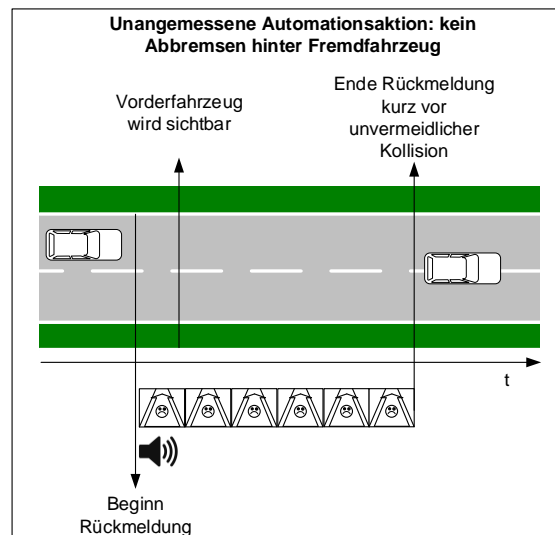
Als weiterer Einflussfaktor soll das zwischenzeitliche Erleben von korrekten Rückmeldungen bzw. positiver Evidenz auf die Entstehung eines Cry Wolf Effektes untersucht werden. Dazu müssen Situationen erlebt werden, in denen eine Rückmeldung gefolgt von einem tatsächlichen Automationsfehler erfolgt. Als Basis für diese Situationen dienen wieder unklare Situationen. In diesen folgt aber nun, nach der Rückmeldung ein Automationsfehler.

Dieser Automationsfehler zeigt sich in der Form, dass das ACC System nicht auf ein vorausfahrendes Fahrzeug reagiert. Um eine Kollision zu vermeiden, muss der Fahrer die Kontrolle übernehmen, indem er selbst die Bremse betätigt. Das Fremdfahrzeug fährt, auch wie die Fremdfahrzeuge in den anderen Situationen, eine Geschwindigkeit von 50 km/h so, dass die Relativgeschwindigkeit bei 50 km/h liegt. Reagiert der Fahrer nicht mit einer Bremsung, vergehen vom Sichtbarwerden des Fremdfahrzeuges aus dem Nebel bis zur Kollision 6,5 Sekunden. Dies ist eine Zeitspanne, in der auch abgelenkte Fahrer reagieren können (Green, 2000). Die Fahrsituationen mit unangemessenem Automationsverhalten werden zur Messung der Stärke eines Cry Wolf Effektes genutzt. Dabei wird die Reaktionszeit des Fahrers auf eine Automationsrückmeldung als Funktion der Stärke des vorher erzeugten Cry Wolf Effektes verstanden (siehe auch Abschnitt „Untersuchungsparadigma“).



**Abbildung 14: „Unklare“ Fahrsituation mit Automationsfehler**

In Situationen mit Automationsfehler hört die Rückmeldung kurz vor einer unvermeidlichen Kollision auf (Abbildung 15). Wenn der Fahrer aufgrund der Rückmeldung rechtzeitig mit einer Bremsung reagiert, hört die Rückmeldung nach einer gewissen Strecke Folgefahrt, mindestens jedoch 100 Meter nach Sichtbarwerden des Fremd-fahrzeuges auf.



**Abbildung 15: Schematische Darstellung des Ablaufs einer Automationsrückmeldung**

Im weiteren Verlauf der Fahrt finden dann wieder zunächst Fahrsituationen statt, die keinen Automationsfehler aber eine Rückmeldung enthalten, um schließlich wieder eine Fahrsituation mit Rückmeldung und Automationsfehler zu erleben (zur exemplarischen Verdeutlichung siehe Abbildung 13 – untere Reihe).

### 5.1.3 Messung des Cry Wolf Effektes

Der Cry Wolf Effekt äußert sich wahrscheinlich in verschiedenen Verhaltensmaßen und kann auf unterschiedliche Weise erfasst werden.

Als Hauptmethode zur Erfassung der Stärke eines Cry Wolf Effektes dient die Erfassung der Fahrerreaktion, genauer der Bremsreaktionszeit in den Situationen mit Rückmeldung und folgendem Automationsfehler.

Solche Situationen können entweder, wie im vorherigen Abschnitt beschrieben, während einer Fahrt erlebt werden, oder erst am Ende einer Fahrt, nachdem vorher je nach Bedingung unterschiedliche häufig falsche Alarme erlebt wurden.

Die Bremsreaktionszeit wird gemessen von Beginn der Rückmeldung bis zum ersten Betätigen des Bremspedals durch den Fahrer.

Da die Bremsreaktionszeit als alleiniges Maß zur Bestimmung der Güte der Fahrerreaktion nicht ausreichen könnte, werden als weitere Maße zur Bestimmung der Güte der Fahrerreaktion die minimale Time to Collision (TTC<sub>min</sub>) und die minimum Distance in Erwägung gezogen. So kann z.B. früh gebremst werden, aber dafür deutlich schwächer, was insgesamt zu einer geringeren minimalen Time to Collision (TTC) und somit einer höheren Kritikalität führen würde. In allen Studien wird sowohl die Bremsreaktionszeit als auch die minimale TTC und die minimum Distance erfasst.

Es zeigt sich in allen drei Teilstudien, dass die Bremsreaktionszeit in sehr hohem Maß sowohl mit der Minimum TTC (-.934 in Studie 3 und -.958 in Studie 1, Korrelationskoeffizient nach Pearson) als auch der minimum Distance (-.882 Pearson in Studie 3) korreliert. Aufgrund der hohen Korrelation wird davon ausgegangen, dass die Bremsreaktionszeit in ausreichendem Maß die Fahrerreaktion abbilden kann. Es wurde daher auf die Darstellung weiterer Maße zur Erschließung der Fahrerreaktion verzichtet.

Eine weitere Möglichkeit zur Erfassung eines Cry Wolf Effektes stellt die Erfassung der Leistung in einer Nebenaufgabe in den Zeiträumen von Beginn einer Rückmeldung bis zu einem bestimmten Punkt im weiteren Verlauf der Fahrsituation dar. Zur Bearbeitung der Nebenaufgabe muss sich der Fahrer visuell von der Fahraufgabe abwenden. Die Fahrerablenkung in diesen Zeitfenstern soll Rückschlüsse auf die Reliance in das Rückmeldesystem und damit auf die Stärke des Cry Wolf Effektes ermöglichen. Zur genaueren Erläuterung der Nebenaufgabe siehe Abschnitt „Fahrerablenkung durch Nebenaufgaben“.

## *5.2 Methodik zur Untersuchung der Effekte auf das Vertrauen in die Automation*

Zur Untersuchung der Auswirkungen unterschiedlicher Rückmeldestrategien auf die allgemeine Erwartung von Automationsfehlern in Form des Vertrauens in die Automation, werden auf Basis des im Abschnitt „Methodik zur Untersuchung der Cry Wolf Anfälligkeit“ vorgestellten Fahrszenarios folgende Bedingungen umgesetzt:

- Zur Untersuchung der Entstehung von Overtrust und als Kontrollbedingung, erleben die Versuchspersonen eine Automation, die in jeder Situation angemessen reagiert und in keiner Situation eine Rückmeldung bzw. Alarm äußert. (Abbildung 16 – erste Reihe)
- Zur Untersuchung der Auswirkungen von Automationsfehlern ohne vorherige Warnung durch ein Rückmeldesystem auf die Entstehung von Misstrust, erleben die Versuchspersonen eine Automation mit Automationsfehlern in der die Automation in keiner Situation vorher eine Rückmeldung bzw. einen Alarm gibt. (Abbildung 16 – vierte Reihe)
- Zur Untersuchung der Auswirkung einer Automationsrückmeldung ohne tatsächlich folgenden Automationsfehler (falscher Alarm) auf das Vertrauen in die Automation, erleben die Versuchspersonen eine Automation, die in jeder Situation angemessen reagiert aber dennoch zwischenzeitlich Automationsrückmeldungen bzw. Alarme gegeben werden. (Abbildung 16 – dritte Reihe)
- Zur Untersuchung der Vermeidung der Entstehung von Misstrust als Effekt eines Rückmeldesystems das vor tatsächlichen Automationsfehlern warnt, erleben die Versuchspersonen eine Automation mit Automationsfehlern vor denen jeweils Automationsrückmeldungen gegeben werden und darüber hinaus in Situationen ohne Fehler auch Automationsrückmeldungen gegeben werden. (Abbildung 16 – vierte Reihe)



**Abbildung 16: Reihe 1: Konditionierung von Overtrust – während der gesamten Fahrt wird kein Automationsfehler und keine Rückmeldung erlebt. Reihe 2: Konditionierung von Misstrust – während der gesamten Fahrt werden mehrere Automationsfehler (rot) aber keine vorherige Automationsrückmeldung erlebt. Reihe 3: Moderierender Einfluss einer Automationsrückmeldung auf die Entstehung von Overtrust. Während der ganzen Fahrt werden Automationsrückmeldungen (gelb) jedoch keine Automationsfehler erlebt. Reihe 4: Moderierender Einfluss einer Automationsrückmeldung auf die Entstehung von Misstrust – vor den erlebten Automationsfehlern wird jeweils eine Automationsrückmeldung (orange) gegeben.**

### 5.2.1 Trustmessung

Die Erfassung des Vertrauens in die Automation soll durch unterschiedliche Messinstrumente ermöglicht werden. Zum einen soll Gesamttrust durch einen Fragebogen, jeweils nach einer Fahrt bzw. einem Fahrabschnitt erfasst werden (siehe Abschnitt Trustfragebögen). Zum anderen soll situatives Vertrauen unmittelbar nach einer Fahrsituation durch eine einfache Einschätzung anhand einer zehnstufigen Skala gemessen werden (siehe Erfassung des situativen Vertrauens).

## 6 Versuchsmaterial & generelles Vorgehen

In diesem Teil werden die über alle drei Studien hinweg verwendeten Versuchsmaterialien vorgestellt.

### 6.1 *Fahrsimulator*

Alle 3 Studien dieser Arbeit werden im selben Fahrsimulator des Instituts für Verkehrssystemtechnik des Deutschen Zentrums für Luft- und Raumfahrt in Braunschweig durchgeführt (Abbildung 17). Der einfache Fahrsimulator besteht aus einem Autositz auf einer Bodenplatte und ist mit einem aktiven Force Feedback Lenkrad sowie einem aktiven Gaspedal und einem passiven Bremspedal ausgestattet (Serienteil zur realistischen Darstellung des Bremspedalwiderstandes). Am Lenkrad befinden sich Lenkstockhebel aus einem Serienfahrzeug. Zur Aktivierung der Automation wird ein Lenkstockhebel verwendet, der in VW Serienfahrzeugen zur Einstellung eines Adaptive Cruise Controls (ACC) genutzt wird. Die Fahrszene wird auf einer 2 x 2 Meter Projektionsfläche in etwa 2 Meter Abstand zum Fahrer dargestellt. Zur linken Seite befindet sich ebenfalls eine 2 x 2 Meter Projektionsfläche auf der die Fahrszene erweitert wurde. Die linke Projektionsfläche ermöglicht einen Schulterblick für Überholmanöver.

Der Geschwindigkeitseindruck wird durch ein Motorengeräusch und eine leichte, von der Motordrehzahl abhängige Vibration der Bodenplatte verstärkt.

Geschwindigkeit, Motordrehzahl als auch der Status der Automation (an, aus) und die Automationsrückmeldung werden für die Studien dieser Arbeit auf einem Display angezeigt, das sich direkt hinter dem Lenkrad befindet. Dabei befindet sich der Tachometer links im Bild, der Drehzahlmesser rechts und die Anzeige des Automationsstatus und der Automationsrückmeldung erfolgt zwischen Drehzahlmesser und Tachometer.

Zur Darstellung einer Nebenaufgabe ist rechts neben dem Lenkrad ein Touchscreen angebracht.



**Abbildung 17: Der für alle drei Studien eingesetzte Fahrsimulator.**

Die im Abschnitt „Methodik zur Untersuchung der Cry Wolf Anfälligkeit“ beschriebene Fahrzeugautomation kann über einen Lenkstockhebel, wie er auch in Serienfahrzeugen zu finden ist, aktiviert werden.

Nach kurzem antippen des Gaspedals beschleunigt das ACC selbständig auf 100 km/h. Um eine standardisierte Geschwindigkeit über alle Versuche und Versuchspersonen zu ermöglichen, kann die Geschwindigkeit durch den Fahrer nicht dauerhaft verstellt werden. Temporäres Übersteuern des ACC durch Betätigen des Gaspedals ist aber möglich. Lässt der Fahrer das Gaspedal los, regelt das ACC wieder auf 100 km/h ein. Durch Betätigen der Bremse wird das ACC in den Standby Modus geschaltet und kann nur durch ein erneutes Betätigen des ACC Lenkstockhebels reaktiviert werden. Die Aktivierung des Systems wird durch ein ACC Symbol, wie es auch in Serienfahrzeugen zu finden ist, zwischen Drehzahlmesser und Tacho angezeigt.

Zur Spurhalteunterstützung wird ein aktives Lenkrad eingesetzt. Kommt der Fahrer von der Spurmitte ab und nähert sich der Fahrbahnmarkierung, wird eine Kraft auf das Lenkrad aufgeschaltet, welche mit geringer werdendem Abstand zur Fahrbahnmarkierung zunimmt, jedoch immer leicht durch den Fahrer übersteuerbar ist.

Zusätzlich zu den Kräften wird eine leichte Vibration auf das Lenkrad aufgeschaltet (Abbildung 9).

Um den Fahrer nicht zu motivieren, sich dauerhaft und vollständig von der Fahraufgabe abzuwenden, sondern, gerade auch im Umfeld des einfachen Fahrsimulators noch ein gewisses Gefährdungsbewusstsein zu erzeugen, ist die Spurhalteunterstützung nicht mittenzentriert ausgelegt. Lenkt der Fahrer gar nicht mit, so kommt es zu einem Aufschaukeleffekt und nach zwei bis dreimaligem Pendeln zwischen den Spurmarkierungen verlässt das Fahrzeug den Fahrstreifen.



Ebenso werden die Fahrer durch Instruktionen darauf hingewiesen, dass sie für die Sicherheit der Fahraufgabe verantwortlich bleiben.

## 6.2 Umsetzung der Rückmeldungssituation

Die Automationsrückmeldung besteht, unabhängig von der Art der Gestaltung, immer aus einem grafischen Symbol und einem akustischen Signal und wird im Display zwischen Drehzahlmesser und Tachometer dargeboten (Abbildung 18).

Ebenfalls unabhängig von der Art der Gestaltung der Automationsrückmeldung ist der zeitliche Ablauf in einer Fahrsituation.

Es ist nicht Ziel der Arbeit herauszufinden wie Unsicherheit am besten zurückgemeldet wird, sondern die Effekte einer funktionierenden Unsicherheitsrückmeldung zu untersuchen!

Daher wird stellvertretend für mögliche Auslegungen einer Unsicherheitsrückmeldung ein Symbol für Automationsunsicherheit ausgewählt und die Bedeutung dieses Symbols den Versuchspersonen vor Fahrtantritt erklärt.



**Abbildung 18: Darstellung der Unsicherheitsrückmeldung im Display zwischen Tachometer und Drehzahlmesser.**

Den Versuchspersonen wird erklärt, dass sich die Automation bei Rückmeldung dieser Automationsunsicherheit unsicher darüber ist, in der gegenwärtigen Situation die richtige Unterstützung aufrechterhalten zu können. Die Fahrer werden ebenso darauf hingewiesen, dass Automationsunsicherheit nichts mit Unsicherheit der Versuchspersonen selbst zu tun hat.

Ähnlich wird verfahren, wenn in einer Teilstudie ein normaler Alarm als Vergleichsbedingung genutzt wird. Es wird ein stellvertretendes Symbol für den Alarm ausgewählt und den Versuchspersonen erklärt, dass die Automation damit vor einem möglichen Automationsfehler warnt. Die verwendeten Symbole sind in den jeweiligen Methodenteilen der Teilstudien dargestellt.

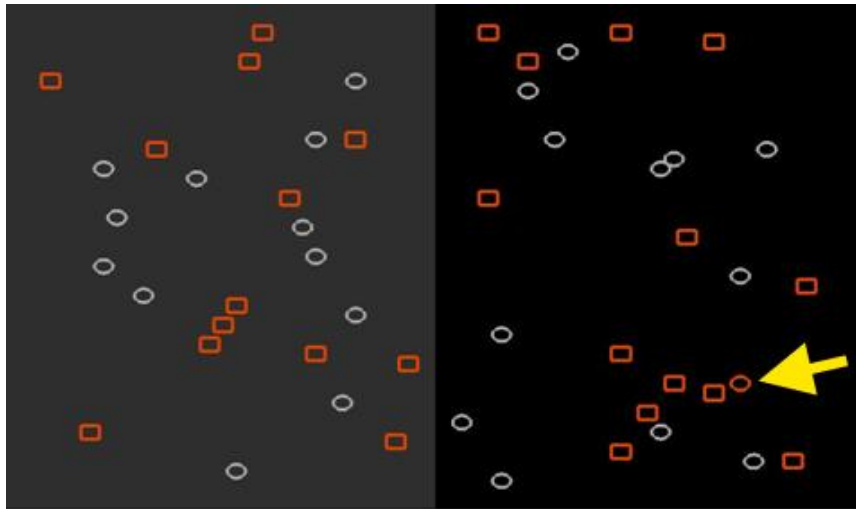
### 6.3 Fahrerablenkung durch Nebenaufgaben

In allen Teilstudien wird, zusätzlich zur Fahraufgabe, eine visuelle Nebenaufgabe angeboten.

Die Nebenaufgabe soll zwei wesentliche Funktionen erfüllen:

- Die Erzeugung von Fahrerablenkung von der primären Fahraufgabe, wie es bei hochautomatisiertem Fahren zu erwarten ist. Somit kann eine „Out of the Loop“ Situation erzeugt werden, anhand der die Wirksamkeit der jeweiligen Rückmeldestrategien, den Fahrer wieder „in the Loop“ zu holen, gemessen werden kann.
- Die Messung von situativem Systemvertrauen über die situationsbezogene Fahrerablenkung. Zu diesem Zweck muss die Nebenaufgabe sensitiv genug sein, um prinzipiell zwischen relativ kurzen Alarmzuständen und entsprechend langen Referenzzuständen ohne Alarm unterscheiden zu können. Die Methode wurde von Rauch, Graddenegger & Krüger (2009) zur Messung von Situationsbewusstsein eingeführt. Es kann angenommen werden, dass die Aufmerksamkeitsverteilung des Fahrers zwischen Fahr- und Nebenaufgabe von der Ausprägung des Vertrauens im Allgemeinen, aber insbesondere auch des situativen Vertrauens in die Automation abhängt.

Bei der eingesetzten Nebenaufgabe handelt es sich um die Surrogate Reference Task (SURT) von Mattes (2003). Die Nebenaufgabe besteht aus einem Touchscreen, der in zwei nebeneinanderliegende Felder unterteilt ist (Abbildung 19). Auf jedem dieser Felder befindet sich eine Anzahl von Ablenkungsobjekten in Form von roten Vierecken und weißen Kreisen. Auf einer der beiden Hälften ist einer der Kreise nicht weiß, sondern rot. Der rote Kreis stellt das Zielobjekt dar. Die Versuchspersonen haben die Aufgabe, die Bildschirmhälfte mit dem Zielobjekt zu identifizieren, und diese per Berührung des Touchscreens auszuwählen. Wird die richtige Hälfte gewählt wird diese grün und rot, wenn die falsche Hälfte gewählt wird. Nach der Identifizierung eines Zielobjektes folgt direkt die nächste Aufgabe mit einer neuen Verteilung der Objekte.



**Abbildung 19: SURT der Pfeil zeigt auf den Zielreiz**

Die Nebenaufgabe ist so leicht gestaltet, dass Versuchspersonen ohne längeres Training sehr schnell eine gute Leistung erbringen können. Durch die Einfachheit der Aufgabe können recht viele Aufgaben in einer recht kurzen Zeit bearbeitet werden, so dass auch innerhalb der recht kurzen Unsicherheits- bzw.

Alarmrückmeldeperioden mehrere Nebenaufgaben bearbeitet werden können.

Dadurch soll die Nebenaufgabe sensitiv genug sein, über die Anzahl der bearbeiteten Nebenaufgaben, den Grad der Fahrerablenkung bzw. des situativen Vertrauens in die Automation zu erfassen.

Um keine übermäßige Beschäftigung mit der Nebenaufgabe und eine völlige Vernachlässigung der Fahraufgabe zu provozieren, werden in der ersten Studie die Versuchspersonen instruiert, dass sie nur in einem solchen Maße Nebenaufgaben bearbeiten sollen wie sie das Fahren als noch sicher erachten. Sie werden weiter instruiert, dass ein sicheres Fahren stets eine höhere Priorität gegenüber der Nebenaufgabe hat und im Zweifel die Nebenaufgabe vernachlässigbar sei. Durch diese Instruktion wird in Kauf genommen, dass die Fahrer in den Testsituationen eher die Aufmerksamkeit auf die Fahraufgabe lenken und somit eher Situationsbewusstsein erlangen können, als sich mit der Nebenaufgabe zu beschäftigen. Jedoch wird es als wichtiger erachtet, dass in der Untersuchungssituation in einem relativ einfachen Fahrsimulator, der Wille zur Vermeidung von Kollisionen nicht eine untergeordnete Rolle spielt.

In den Studien zwei und drei wird dieser Trade-Off zwischen Nebenaufgabenbearbeitung und der Notwendigkeit Kollisionen zu vermeiden, weiter intensiviert und mit einem Belohnungs-Bestrafungssparadigma gearbeitet.

Die Versuchspersonen werden hier in eine Wettbewerbssituation untereinander versetzt. Sie werden instruiert, dass diejenige Versuchsperson mit der besten Nebenaufgabenleistung einen Preis in der Höhe von etwa 10 bis 15 Euro erhält.

Außerdem wird in Aussicht gestellt, dass sie ihre Leistung im Gesamtversuchspersonenranking ablesen können sobald die Studie abgeschlossen ist.

Eine gute Nebenaufgabenleistung ist dabei definiert, als das Verhältnis der insgesamt bearbeiteten Aufgaben zu den falsch bearbeiteten Aufgaben.

Um auf der anderen Seite eine starke Motivation zu erzeugen, Kollisionen und insgesamt sicherheitskritische Events (Überfahren der Spurmarkierung etc.) zu vermeiden, werden die Versuchspersonen instruiert, dass sie automatisch dann von dem Wettbewerb disqualifiziert werden, wenn unabsichtlich die Spur verlassen, die Leitplanke touchiert wird oder seitliche und frontale Kollisionen mit anderen Fahrzeugen stattfinden.

Die Versuchspersonen müssen also einen Trade-Off zwischen Nebenaufgabenleistung und Fahrleistung bilden, der es ermöglicht einerseits die Nebenaufgabenleistung zu maximieren und andererseits die Fahraufgabe und Überwachung des Systems nicht übermäßig zu vernachlässigen, um einen Totalverlust zu vermeiden.

## *6.4 Fragebögen und weitere Messinstrumente*

### 6.4.1 Trustfragebögen

Da es zu Beginn der Studie 1, 2009, im gesamten deutschsprachigen Raum, keinen dem Autor bekannten Fragebogen zur Erfassung des Vertrauens in die Automation gab, wurde zur Erfassung des Vertrauens in die Automation ein 5 Items umfassender englischsprachiger Fragebogen, der von Master et. Al. (2000) entwickelt wurde, übersetzt. Die 5 erfassten Items des Fragebogens basieren auf einem von Muir (1994) vorgeschlagenen Trust Modell und von Barber (1983) und Rempel et al. (1985) vorgeschlagenen Trust Dimensionen.

Die umfassten Dimensionen sind Kompetenz, Zuverlässigkeit, Vorhersagbarkeit, Vertrauen und als generelle Trust Dimension das allgemeine Vertrauen. Die 5 Trustdimensionen sind durch folgende Fragen definiert:

- Kompetenz: In welchem Ausmaß erfüllt die Automation die Fahraufgabe effektiv?
- Vorhersagbarkeit: In welchem Ausmaß können Sie das Verhalten der Automation mit einem Mindestgrad an Sicherheit vorhersagen?
- Zuverlässigkeit: In welchem Ausmaß ist die Automation fehlerfrei?

- Vertrauen: In welchem Ausmaß haben Sie Vertrauen darin, dass die Automation bestimmte Aufgaben effektiv bearbeitet, ohne dass Sie das prüfen können?
- Allgemeines Vertrauen: In welchem Ausmaß vertrauen Sie der Automation im Allgemeinen?

Die Beantwortung aller Trustdimensionen erfolgt anhand einer 10fach abgestuften Skala zwischen den Polen „Überhaupt nicht“ und „Sehr hoch“. Die Versuchspersonen sollen den zu ihrer Einschätzung passenden Abschnitt auf der Skala mit einem Kreuz markieren.

Überhaupt nicht |-----| Sehr hoch

Im Laufe der Durchführung der Studien zu dieser Arbeit, wurde ein deutschsprachiger und validierter Trustfragebogen von Wiczorek (2011) verfügbar. Dieser entspricht in den wesentlichen Dimensionen dem bereits eingesetzten Instrument. Jede einzelne Dimension ist jedoch durch mehrere Items operationalisiert. Der Fragebogen wurde, neben dem bereits eingesetzten Fragebogen in Studie 2 verwendet. Da sich aber zeigte, dass sich die Ergebnisse beider Fragebögen nicht differenzierten, werden nur die Ergebnisse des übersetzten Fragebogens berichtet.

#### 6.4.2 Erfassung des situativen Vertrauens

Das situative Vertrauen in die Automation wird in den drei Teilstudien nach jeder Fahrsituation, unmittelbar nachdem das Fremdfahrzeug nicht mehr sichtbar und die Situation „aufgelöst“ ist, erfasst. Dazu wird den Versuchspersonen durch den Versuchsleiter die Frage: „In welchem Ausmaß haben Sie der Automation in der gerade erlebten Situation insgesamt vertraut?“ gestellt. Die Frage soll anhand einer 10 stufigen Skala beantwortet werden, in der 1 = gar nicht vertraut und 10 = absolut vertraut bedeutet. Die Skala wird den Versuchspersonen vor Fahrtantritt erklärt und ist während der Fahrt einsehbar.

#### 6.4.3 Erfassung der Akzeptanz in die Automation

Die Akzeptanz in die Automation wird nach einer Fahrt, bzw. nach einem Fahrtabschnitt, durch die Abfrage der Nutzungstendenz erfasst. Versuchspersonen sollen dafür die Frage, „Würden Sie die Automation ausschalten“ auf einer 7stufigen Skala, die von 1 = „Ja, definitiv ausschalten“ bis 7 = „Nein, definitiv angeschaltet lassen“ reichte, beantworten.

#### 6.4.4 Erfassung der Kooperation

In Studie 1 wird neben der Akzeptanz der Automation auch die subjektiv empfundene Kooperation mit der Automation abgefragt. Dafür wird ein Fragebogen, der von Skjerve (2002) ursprünglich zur Erfassung der subjektiv empfundenen Kooperation zwischen Mensch und Maschine im Kontext der Überwachung von Atomkraftwerksleitwarten entwickelt wurde, genutzt und zur Verwendung im Kontext assistiertes / automatisiertes Fahren angepasst. Der Fragebogen besteht insgesamt aus 6 Items, welche die subjektive Kooperationseinschätzung der Versuchspersonen anhand einer 7 stufigen Skala erfassen:

1. Ob relevante Informationen durch das System dargeboten wurden
2. Ob diese Informationen rechtzeitig dargeboten wurden
3. Ob diese Informationen verstanden wurden
4. Das Ausmaß der gewünschten Automationsaktivitäten
5. Das Ausmaß der erwarteten Automationsaktivitäten
6. Die Einschätzung der Kooperation insgesamt

#### 6.5 Generelles Vorgehen: Stichproben

Die Versuchspersonen für die drei Teilstudien wurden sämtlich aus dem Probandenpool des DLR Instituts für Verkehrssystemtechnik akquiriert. Es wurde bei den Studien 2 und 3 darauf geachtet, dass Versuchspersonen nicht bereits vorher an einer vorherigen Teilstudie teilgenommen hatten. Dadurch soll Sekundärvarianz durch Vorwissen eliminiert werden. Aus versuchsökonomischen Gründen, wegen einer limitierten Anzahl an verfügbaren Versuchspersonen und der Bedingungsanzahl in den Teilstudien, wird die Versuchspersonenzahl pro Bedingung eingeschränkt und von einer Randomisierung der Bedingungs-zuteilung abgesehen. Die Versuchspersonen werden möglichst gleichmäßig nach Alter und Geschlecht auf die Subgruppen parallelisiert.

Es wurden nach der Datenerhebung diese Versuchspersonen von der Datenanalyse ausgeschlossen, bei denen die Bedingungsmanipulation des Faktors niedriges bzw. hohes Cry Wolf Potenzial nicht funktionierte. So reagierten einige der Fahrer in den Rückmeldungssituationen ohne Automationsfehler trotzdem mit einem Eingriff, indem sie z.B. bremsen, obwohl die Automation rechtzeitig selbst gebremst hätte. Bliss, Gilson & Deaton (1995) beobachteten, dass etwa 10% aller ihrer Versuchspersonen zu solchen „Extreme Respondern“ gehörten und nach einer „Alles oder Nichts“ Strategie vorgehen. Diese Versuchspersonen konnten durch diese frühzeitige, vorbeugende Reaktion natürlich nicht mitbekommen, dass die Automation in dieser Situation eigentlich richtig reagiert hätte. Für diese Versuchspersonen ist die Rückmeldung in diesem Sinne immer „korrekt“. Die ursprünglich angestrebte

Häufigkeit von erlebten Rückmeldungen ohne Automationsfehler kann für diese Versuchspersonen also nicht erreicht werden. Ein Ausschluss erfolgte dann, wenn Versuchspersonen in mehr als einem Fall in der Situationsart „unklar – Fahrzeug halb auf eigener Spur“ bremsen.

Da der Ausschluss erst nachträglich erfolgte, konnten die Parallelisierungskriterien nicht immer exakt eingehalten werden.

In den jeweiligen Abschnitten zu den Teilstudienergebnissen wird berichtet, wie viele Versuchspersonen letztendlich in die Analyse einfließen.

## ***6.6 Generelles Vorgehen: Datenauswertung***

### **6.6.1 Varianzanalyse & eingesetzte Korrekturverfahren**

Insgesamt wird zur Auswertung der Daten das Verfahren der Varianzanalyse, sowohl einfaktoriell, mehrfaktoriell als auch für Messwiederholungen angewandt.

Als Voraussetzung für die Durchführung einer Varianzanalyse mit Messwiederholung gilt, dass die Varianzen der Mittelwertdifferenzen zwischen den Messzeitpunkten homogen sind (Sphärizität). Es wurde bei jeder Varianzanalyse mit Messwiederholung auf Sphärizität getestet. Wenn Sphärizität nicht angenommen werden kann, wird, nach der Empfehlung von Girden (1992), die Huynh-Feldt Korrektur angewandt, wenn Epsilon  $>.75$  ist, da die Greenhouse-Geisser Korrektur bei Epsilon-Werten  $>.75$  als zu konservativ gilt. Bei Epsilon Werten  $<.75$  wird entsprechend nach Greenhouse-Geisser korrigiert. Zu den einzelnen Ergebnissen wird nicht berichtet ob, bzw. welches Korrekturverfahren angewandt wird, wenn Korrekturverfahren verwendet werden, werden die korrigierten „Degrees of Freedom“ berichtet.

### **6.6.2 Non-Parametrische Tests**

Da in den einzelnen Teilstudien nur geringe Stichprobengrößen vorhanden sind, kann nicht immer von einer Normalverteilung der Stichproben ausgegangen werden, es werden daher zusätzlich zur Varianzanalyse jeweils non-parametrische Tests durchgeführt.

Bei Vergleichen von je zwei voneinander unabhängigen Verteilungen wird der Mann-Whitney-U Tests genutzt, bei Wiederholungsmessungen der Friedmann-Test.

Die Ergebnisse der non-parametrischen Tests werden nur dann berichtet, wenn das Ergebnis des Tests im Widerspruch zu den Ergebnissen der ANOVA steht.

## 7 Studie 1

### 7.1 *Ziele Studie 1*

Die erste durchgeführte Studie soll zunächst eine grobe Orientierung im Feld der Unsicherheitsrückmeldung ermöglichen und einige notwendige Kriterien zur weiteren Betrachtung in Folgestudien überprüfen. Kurz zusammengefasst sind folgende Punkte von Interesse:

- Die Effektivität der Wirkung der Unsicherheitsrückmeldung auf die Kontrollierbarkeit.
- Erste Aussagen zur Anfälligkeit der Unsicherheitsrückmeldung gegenüber einem Cry Wolf Effekt.
- Die Wirkung auf subjektiv empfundenes Gesamtvertrauen in die Automation.
- Die Auswirkung auf die subjektiv empfundene Kooperation und die Akzeptanz in die Automation.

Die einzelnen Studienziele und die mit diesen Zielen verbundenen Forschungsfragen und Hypothesen werden im folgenden Teil erläutert:

**Ziel 1:** Betrachtung der generellen Effektivität einer Unsicherheitsrückmeldung. Wesentliches notwendiges Kriterium zur weiteren Betrachtung von unsicherheitsbasierten Rückmeldungen ist der Nachweis einer grundsätzlichen Effektivität hinsichtlich der Verbesserung der Kontrollierbarkeit eines Automationsfehlers. Eine Unsicherheitsrückmeldung muss also zu einer besseren Kontrollierbarkeit führen als eine Vergleichsbedingung ohne Rückmeldung.

**Forschungsfrage:** Kann durch eine unverbindliche Rückmeldung wie die Unsicherheitsrückmeldung die Kontrollierbarkeit eines Automationsfehlers verbessert werden?

**Hypothese:** Auch eine unverbindliche Rückmeldung, wie die Unsicherheitsrückmeldung, führt zu einer besseren Kontrollierbarkeit eines Automationsfehlers verglichen mit keiner Rückmeldung.

**Ziel 2:** Untersuchung der Anfälligkeit einer Unsicherheitsrückmeldung für einen Cry Wolf Effekt durch Vergleich der Effektivität der Unsicherheitsrückmeldung bei keiner Erfahrung von „falschen Alarmen“ und bei Erfahrung vieler „falscher Alarme“. Sowie Überprüfung der Kontrollannahme, dass die Erfahrung einer fehlerfreien Automation zu einer Verschlechterung der Kontrollierbarkeit eines Automationsfehlers führt.

**Forschungsfrage:** Führt die wiederholte Erfahrung einer Unsicherheitsrückmeldung mit Ausbleiben eines Automationsfehlers zu einer



Verringerung der Effektivität der Unsicherheitsrückmeldung in einer Kontrollierbarkeitssituation?

**Hypothese:** Wird in einer Fahrt erst eine Reihe von Unsicherheitsrückmeldungen ohne nachfolgenden Automationsfehler erlebt und dann eine Rückmeldung mit Automationsfehler, so ist die Kontrollierbarkeit des Automationsfehlers nicht schlechter, als in einer Vergleichsbedingung mit Unsicherheitsrückmeldung in der der Automationsfehler bereits zu Beginn einer Fahrt passiert und zuvor keine Bedingung mit Rückmeldung ohne Fehler erlebt wurde.

**Hypothese:** Die Kontrollierbarkeit des Automationsfehlers ist in der Bedingung Unsicherheitsrückmeldung, auch in dem spät erlebten Automationsfehler, besser als in einer Vergleichsbedingung ohne Rückmeldung.

**Kontrollannahme:** Führt die Erfahrung einer fehlerfreien Automation zu einer Verschlechterung der Kontrollierbarkeit in einer Bedingung ohne Rückmeldung?

**Hypothese:** In einer Bedingung ohne Rückmeldung ist die Kontrollierbarkeit eines Automationsfehlers besser, wenn dieser direkt zu Anfang einer Fahrt erlebt wird und noch keine Erfahrung einer fehlerfreien Automation gemacht wurde, verglichen mit der Kontrollierbarkeit eines Automationsfehlers, wenn dieser erst im späteren Verlauf einer Fahrt auftritt und bis dahin eine fehlerfreie Automation erlebt wurde.

**Ziel 3:** Untersuchung der Auswirkungen des Erlebens von Automationsfehlern auf die generelle Erwartung weiterer Automationsfehler und Untersuchung der Wechselwirkung von Fehlererfahrung und Unsicherheitsrückmeldung auf die Kontrollierbarkeit von Automationsfehlern.

**Forschungsfrage:** Wirken situative Erhöhung der Fehlererwartung durch Unsicherheitsrückmeldung und allgemeine Erhöhung der Fehlererwartung additiv?

**Hypothese:** Die Erhöhung der Fehlererwartung durch Fehlererfahrung und situative Erhöhung der Fehlererwartung durch Automationsunsicherheit wirken additiv. D.h. die Kontrollierbarkeit eines Automationsfehlers wird besser, wenn Fehlererfahrung und Unsicherheitsrückmeldung vorliegen. Innerhalb der Unsicherheitsbedingung ist die Kontrollierbarkeit beim ersten erlebten Automationsfehler schlechter als beim zweiten bzw. folgenden Automationsfehler. Die Verbesserung ist unabhängig von der Anzahl der vorher erlebten Rückmeldungen ohne Automationsfehler und kann somit nicht durch einen Cry Wolf Effekt beim ersten Automationsfehler erklärt werden.

**Hypothese:** In einer Bedingung ohne Rückmeldung erhöht die Erfahrung eines Automationsfehlers die Erwartung weiterer Fehler, daher werden folgende Automationsfehler besser kontrolliert. Jedoch bleibt die Kontrollierbarkeit in allen Automationsfehlerbedingungen schlechter als in der Vergleichsbedingung mit Unsicherheitsrückmeldung. Die Differenz zur Bedingung mit Unsicherheitsrückmeldung bleibt in allen Fehlersituationen gleich und stellt den Effekt der situativen Fehlererwartung durch Rückmeldung dar.

**Ziel 4:** Untersuchung der Auswirkungen einer Unsicherheitsrückmeldung auf das Vertrauen in die Automation im Vergleich zu keiner Rückmeldung. Durch diese Betrachtung soll beurteilt werden können, ob eine Unsicherheitsrückmeldung durch die Art der Rückmeldung eines Automationszustandes als Automationsunsicherheit ein anderes Vertrauen in die Automation, welches die Möglichkeit von Automationsfehlern einschließt, erzeugen kann.

**Forschungsfrage:** kann durch eine Unsicherheitsrückmeldung das Vertrauen in die Automation angepasst werden in dem Sinne, dass durch die Rückmeldung von Automationsunsicherheit Übervertrauen und Misstrauen vermieden wird?

**Hypothese:** Das Erleben einer fehlerfreien Automation führt in einer Bedingung ohne Rückmeldung zu einem hohen Vertrauen in die Automation. Werden während dieser Fahrt aber Unsicherheitsrückmeldungen ohne Automationsfehler erlebt, so ist das Vertrauen in dieser Bedingung niedriger als in der Bedingung ohne Rückmeldung, obwohl kein echter Automationsfehler erlebt wird.

**Hypothese:** Die Rückmeldung von Automationsunsicherheit führt zu einer besseren situativen Erwartung eines Automationsfehlers, was sich puffernd auf die Tendenz zur Entwicklung eines Misstrauens in die Automation auswirkt. Wird vor Automationsfehlern Unsicherheit zurückgemeldet, dann ist das Gesamtvertrauen in der Unsicherheitsbedingung höher als in einer Vergleichsbedingung, in der Automationsfehler ohne Rückmeldung erlebt werden.

**Ziel 5:** Untersuchung der Auswirkung der Unsicherheitsrückmeldung auf die subjektiv wahrgenommene Kooperation mit der Automation und die Akzeptanz des Gesamtsystems.

**Forschungsfrage:** In wie weit führt eine Rückmeldung, auch wenn auf diese nicht immer Automationsfehler folgen, zu einer besseren Akzeptanz einer Automation die Fehler macht?

Diesbezüglich wurde keine Hypothese formuliert. Es sollte stattdessen getestet werden, ob es insgesamt einen Unterschied zwischen keiner Rückmeldung und Unsicherheitsrückmeldung hinsichtlich der Akzeptanz der Automation insgesamt gibt, und falls ja, in welche Richtung dieser geht.

## 7.2 Methode Studie 1

Die Studie wird in dem im Gesamtmethodenteil beschriebenen Fahrsimulator durchgeführt. Für die Umsetzung der in Studie 1 vorgenommen Fahrten im Simulator, wird auf das im Gesamtmethodenteil beschriebene Grundfahrtszenario zurückgegriffen. Das heißt, Versuchspersonen fahren, in einem Fahrzeug mit einer integrierten Längs- und Querführung konstant 100 km/h, über eine nebelige zweispurige und gerade Autobahn. Während dieser Fahrt erleben die Versuchspersonen unterschiedliche Situationen mit einem Fremdfahrzeug. Diese sind „klare“ Situationen, in denen nie eine Rückmeldung erfolgt und die Automation keinen Fehler macht, und unklare Situationen.

In den unklaren Situationen können, je nach Bedingung, Unsicherheitsrückmeldungen gegeben werden, denen jedoch nicht immer ein Automationsfehler folgen muss. Während der Fahrt soll eine Nebenaufgabe bearbeitet werden.

Es gibt folgende klare Situationen:

- Vorbeifahrt an einem Fremdfahrzeug das deutlich auf dem rechten Fahrstreifen fährt (Abbildung 20 links)
- Abbremsen und Folgefahrt hinter einem langsameren Fremdfahrzeug das deutlich auf dem linken Fahrstreifen fährt (Abbildung 20 rechts).

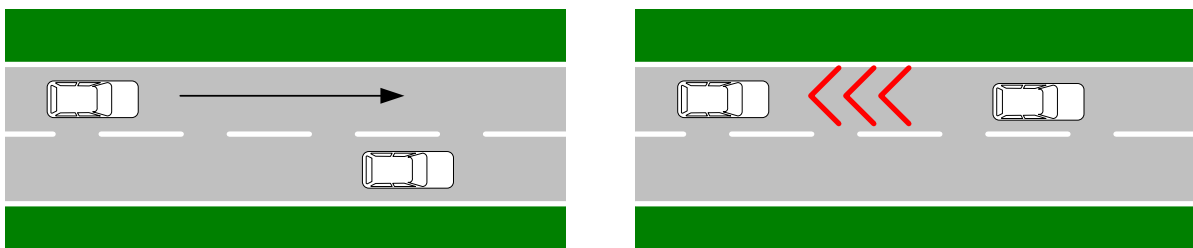
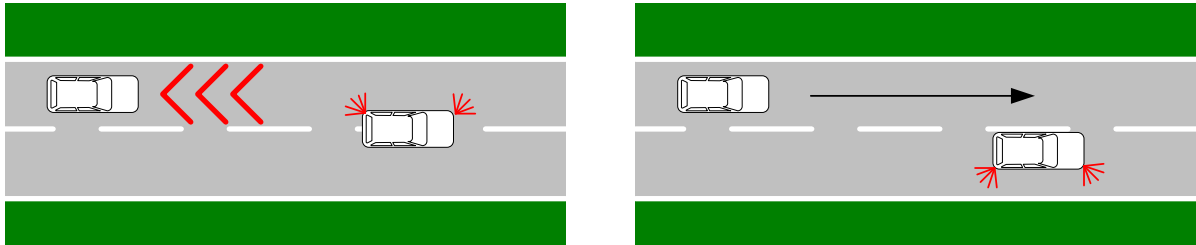


Abbildung 20: „Klare“ Fahrsituationen.

Insgesamt gibt es zwei Typen von unklaren Situationen ohne Automationsfehler:

- Abbremsen und Folgefahrt hinter einem Fahrzeug, das nicht eindeutig einem Fahrstreifen zuzuordnen ist und deutlich in den Fahrstreifen des Ego-Fahrzeuges hereinragt (Abbildung 11 links).

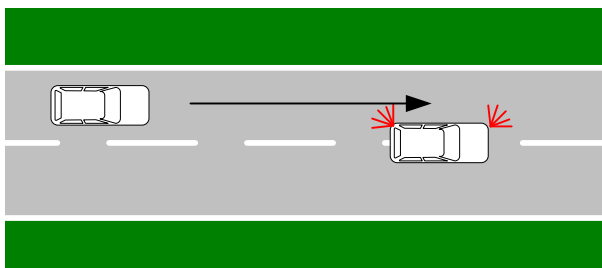
- Vorbeifahrt an einem Fremdfahrzeug, das sich noch ganz auf dem rechten Fahrstreifen befindet und sehr nah an der Spurmarkierung fährt (Abbildung 11 rechts)



**Abbildung 21: „Unklare“ Fahrsituationen ohne Automationsfehler.**

Es gibt einen Typ „unklare Situation“ mit einem Automationsfehler (Abbildung 14).

- Hier bremsst die Automation hinter einem nach rechts versetzt fahrendem Fahrzeug nicht ab, der Fahrer muss mit einer Bremsung eingreifen.



**Abbildung 22: „Unklare“ Fahrsituation mit Automationsfehler**

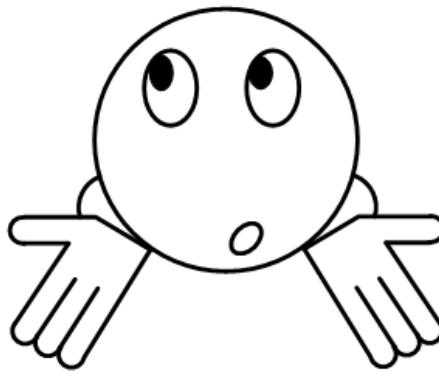
## 7.2.1 Versuchsdesign je Teilziel:

### 7.2.1.1 Ziel 1 Design: Faktor Rückmeldung

Zur Ermöglichung des Vergleichs zwischen Unsicherheitsrückmeldung und keiner Rückmeldung wird ein zweistufiger Zwischengruppenfaktor definiert:

1. Stufe 1: Unsicherheitsrückmeldung in unklaren Situationen
2. Stufe 2: Kontrollgruppe – keine Rückmeldung in unklaren Situationen

Die Unsicherheitsrückmeldung wird durch das in Abbildung 23 abgebildete Symbol im Display hinter dem Lenkrad angezeigt. Das Erscheinen wird durch einen Ton begleitet, der auch bei visueller Ablenkung das Vorhandensein einer Rückmeldung signalisieren soll. Als prototypisches Symbol für Automationsunsicherheit wird ein Emoticon gewählt, das Unsicherheit ausdrücken soll. Die Bedeutung des Emoticons wird den Versuchspersonen nach dem im Gesamtmethodenteil beschriebenen Vorgehen instruiert. Der zeitliche Ablauf der Unsicherheitsrückmeldung entspricht ebenso dem im Gesamtmethodenteil beschriebenen Schema.



**Abbildung 23: Abbildung des Unsicherheitssymbols, welches in der Bedingung Unsicherheitsrückmeldung in unklaren Situationen zurückgemeldet wurde.**

**Abhängige Variable(n):** Kontrollierbarkeit (Reaktionshäufigkeit & Reaktionszeit) der Automationsfehler in unklaren Situationen.

#### **7.2.1.2 Ziel 2 Design: Faktor Fehlerfreiheit - Zeitpunkt erster Automationsfehler**

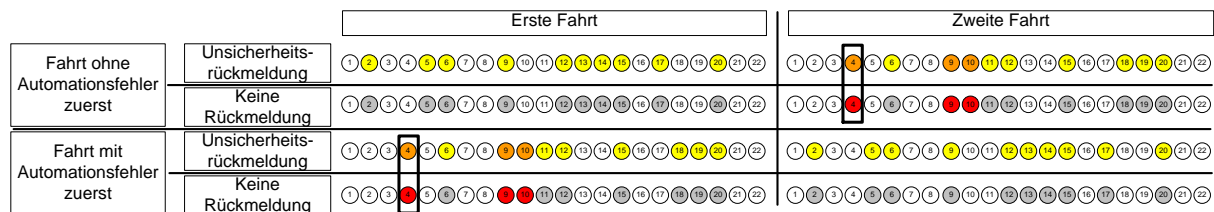
Ergänzend zum Faktor Rückmeldung wird zur Erfassung eines möglichen Cry Wolf Effektes und der Auswirkungen auf das Vertrauen in die Automation der Wiederholungsfaktor Zeitpunkt der Fehlererfahrung eingeführt. Dieser hat zwei Stufen:

1. Stufe: Frühe Fehlererfahrung, Automationsfehler wird früh in erster Fahrt erlebt
2. Stufe: Späte Fehlererfahrung, Automationsfehler wird erst in zweiter Fahrt erlebt

Die Versuchspersonen unternehmen zwei Fahrtabschnitte mit je 22 Fahrsituationen, die sich aus den Situationen der Abbildungen 1 – 3 zusammensetzten (Abbildung 24). Um den Einfluss von Reihenfolgeeffekten zu minimieren, werden die klaren und unklaren Situationen in den zwei Fahrten unterschiedlich verteilt.

In der Stufe frühe Fehlererfahrung absolvierten die Versuchspersonen die erste Fahrt mit Automationsfehlern (Abbildung 24 untere zwei Reihen). Der Automationsfehler passiert direkt in der ersten unklaren Situation. In der Stufe späte Fehlererfahrung reagiert die Automation in den unklaren Situationen der ersten Fahrt fehlerfrei (Abbildung 24 – obere zwei Reihen). In der zweiten Fahrt werden dann Automationsfehlern erlebt.

In Kombination mit dem Faktor Rückmeldung ergab sich ein 2 x 2 faktorielles Mixed Models Design (Abbildung 24).



**Abbildung 24:** zeigt eine Übersicht der Fahrsituationen in den unterschiedlichen Fahrbedingungen und Fahrtabschnitten. Weiße Kreise stehen für klare Fahrsituationen ohne Rückmeldung und ohne Automationsfehler; graue Kreise stehen für unklare Situationen ohne Rückmeldung und ohne Automationsfehler; gelbe Kreise stehen für unklare Situationen mit Rückmeldung aber ohne Automationsfehler; orangene Kreise stehen für Situationen mit Rückmeldung und Automationsfehler; rote Kreise stehen für Situationen ohne Rückmeldung mit Automationsfehler

**Abhängige Variable(n):** Kontrollierbarkeit (Reaktionshäufigkeit & Reaktionszeit) des jeweils ersten erlebten Automationsfehlers in unklaren Situationen – in Abbildung 24 markiert mit schwarzen Rahmen.

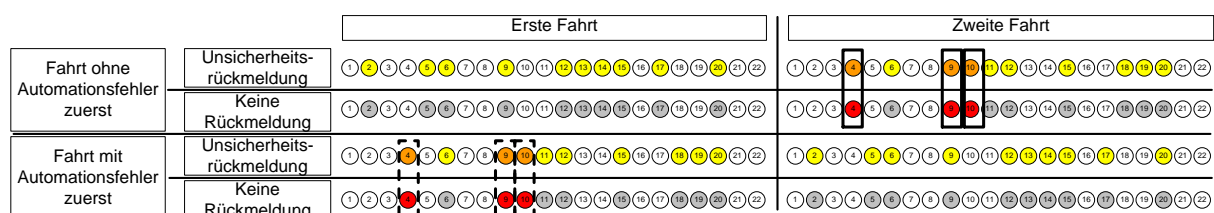
### 7.2.1.3 Ziel 3 Design: Faktor Fehlererfahrung

Zur Untersuchung des Einflusses des Erlebens mehrerer Automationsfehler auf die Erwartung und Kontrollierbarkeit späterer Automationsfehler und zur Untersuchung der Wechselwirkung dieser Fehlererfahrung mit der An- bzw. Abwesenheit einer Unsicherheitsrückmeldung, werden in den Fahrtabschnitten mit Automationsfehlern 3 von diesen hintereinander erlebt (Abbildung 25 untere zwei Reihen, orange und rot markierte Situationen). Entsprechend ergeben sich 3 Faktorstufen des Faktors Fehlererfahrung:

1. Keine Fehlererfahrung
2. Eine vorherige Fehlererfahrung
3. Zwei vorherige Fehlererfahrungen

In Kombination mit dem Faktor Rückmeldung resultiert ein 2 x 3 faktorielles Mixed Models Design.

Ebenso wird betrachtet, ob der Zeitpunkt des Erlebens der Automationsfehler, entweder erst in der zweiten Fahrt oder schon in der ersten Fahrt einen Effekt hat (Abbildung 25 – gestrichelte Kästen vs. durchgehende Kästen).



**Abbildung 25:** Übersicht der Fahrsituationen mit Automationsfehler (rot bzw. orange)

**Abhängige Variable(n):** Kontrollierbarkeit (Reaktionshäufigkeit & Reaktionszeit) von drei aufeinander folgenden Automationsfehlern in unklaren Situationen – in Abbildung 25 markiert mit schwarzen Rahmen.

#### 7.2.1.4 Ziel 4 Design: Vertrauen in die Automation

Zur Untersuchung des Vertrauens in die Automation wird nach jeder Fahrt eine Pause gemacht in denen die Versuchspersonen Fragebögen zum Vertrauen in die Automation bearbeiten (Abbildung 26).

**Abhängige Variable:** Gesamtvertrauen (siehe Abschnitt: Trustfragebögen).

Reihenfolge der Fahrten mit und ohne Automationsfehler					
		Erste Fahrt		Zweite Fahrt	
Fahrt ohne Automationsfehler zuerst	Unsicherheitsrückmeldung	Fahrt ohne Fehler	Fragebogen Gesamttrust	Fahrt mit Fehler	Fragebogen Gesamttrust
	Keine Rückmeldung	Fahrt ohne Fehler	Fragebogen Gesamttrust	Fahrt mit Fehler	Fragebogen Gesamttrust
Fahrt mit Automationsfehler zuerst	Unsicherheitsrückmeldung	Fahrt mit Fehler	Fragebogen Gesamttrust	Fahrt ohne Fehler	Fragebogen Gesamttrust
	Keine Rückmeldung	Fahrt mit Fehler	Fragebogen Gesamttrust	Fahrt ohne Fehler	Fragebogen Gesamttrust

**Abbildung 26: schematischer Aufbau der Untersuchung mit unterschiedlichen Versuchsgruppen, Reihenfolge der Fahrten und den Fragebogensituationen.**

#### 7.2.1.5 Ziel 5 Design: Akzeptanz

Zur Erfassung der Auswirkungen auf die Akzeptanz bearbeiten die Versuchspersonen nach Beendigung der zweiten Fahrt einen Kooperations- und Akzeptanzfragebogen (siehe Abschnitt: Erfassung der Akzeptanz in die Automation).

## 7.2.2 Gesamtversuchsschema

Alle Versuchsdesigns der Teilziele zusammengefasst ergaben folgendes übergeordnetes Versuchsschema (Tabelle 4):

**Tabelle 4: Kombination der Faktoren / Versuchsplan**

				Faktor Reihenfolge der Fahrttypen (Within)	
				Fahrt 1	Fahrt 2
Faktor Rückmeldung (Between)	Unsicherheitsrückmeldung	Faktor: Zeitpunkt des ersten Automationsfehlers / Reihenfolge der Fahrten (Between / Within)	Bed Unreliable Automation in erster Fahrt	<b>Automation mit Fehlern</b> – Automation macht <u>Fehler</u> in einigen unklaren Situationen In <u>allen</u> unklaren Situationen wird Unsicherheit zurückgemeldet (Innergruppenfaktor Fehlererfahrung innerhalb dieser Fahrt)	<b>Fehlerfreie Automation</b> – Die Automation macht <u>keine Fehler</u> in den unklaren Situationen- In <u>allen</u> unklaren Situationen wird Unsicherheit zurückgemeldet
			Bed Reliable Automation in erster Fahrt	<b>Fehlerfreie Automation</b> – Die Automation macht <u>keine Fehler</u> in den unklaren Situationen- In <u>allen</u> unklaren Situationen wird Unsicherheit zurückgemeldet	<b>Automation mit Fehlern</b> – Automation macht <u>Fehler</u> in einigen unklaren Situationen in <u>allen</u> unklaren Situationen wird Unsicherheit zurückgemeldet (Innergruppenfaktor Fehlererfahrung innerhalb dieser Fahrt)
	„Keine Rückmeldung“	Faktor: Zeitpunkt des ersten Automationsfehlers / Reihenfolge der Fahrten (Between / Within)	Bed Unreliable Automation in erster Fahrt	<b>Automation mit Fehlern</b> – Automation macht Fehler in einigen unklaren Situationen In <u>keiner</u> unklaren Situation wird etwas zurückgemeldet (Innergruppenfaktor Fehlererfahrung innerhalb dieser Fahrt)	<b>Fehlerfreie Automation</b> – Die Automation macht keine Fehler in den unklaren Situationen In <u>keiner</u> unklaren Situation wird etwas zurückgemeldet
			Bed Reliable Automation in erster Fahrt	<b>Fehlerfreie Automation</b> – Die Automation macht keine Fehler in den unklaren Situationen In <u>keiner</u> unklaren Situation wird etwas zurückgemeldet	<b>Automation mit Fehlern</b> – Automation macht Fehler in einigen unklaren Situationen In <u>keiner</u> unklaren Situation wird etwas zurückgemeldet (Innergruppenfaktor Fehlererfahrung innerhalb dieser Fahrt)

## 7.2.3 Vorgehen

Nach einer Begrüßung durch den Versuchsleiter wurden den Versuchspersonen ein Aufklärungsbogen und eine Einverständniserklärung über die Aufzeichnung von Daten sowie ein demographischer Fragebogen vorgelegt. Danach wurden sie in das Thema der Fahrerassistenz und des teil- und hochautomatisierten Fahrens eingeführt, indem beispielhafte Assistenzsysteme, insbesondere ein Adaptive Cruise Control (ACC) und ein Lane-Keeping System (HC), bzw. die Kombination aus beiden, beschrieben wurden. Dies wurde getan, da die meisten Versuchspersonen kein Wissen über bereits auf dem Markt befindliche Assistenz und Automationssysteme und somit kaum Vorstellungen bezüglich des Themas Automatisiertes Fahren hatten. Auf diese Weise konnten die Versuchspersonen



„abgeholt“ werden, um sich so besser in die spätere Fahraufgabe einfinden zu können.

Anschließend wurden die Versuchspersonen, nach einem Parallelisierungsschema Aufgrund Alter und Geschlecht, einer der vier Experimentalgruppen zugeordnet. Daraufgehend, wurde den Versuchspersonen der Fahrsimulator erklärt und sie bekamen Gelegenheit, für etwa 15 Minuten mit dem System zu trainieren. Das Training diente dazu, die Versuchspersonen mit dem Fahren im Fahrsimulator und dem Fahren mit dem Assistenzsystem vertraut zu machen. Im Training wurde auch die Nebenaufgabe eingeführt und den Versuchspersonen wurde Gelegenheit gegeben, die letzten fünf Minuten des Trainings das Bearbeiten der Nebenaufgabe während der Fahrt zu trainieren.

Zur Vermeidung von Sekundärvarianz durch unterschiedliche Interpretation der Unsicherheitsrückmeldung, wurde nach dem Training der Unsicherheitsgruppe die Unsicherheitsrückmeldung demonstriert und erklärt (siehe Umsetzung der Rückmeldungssituation). Den Versuchspersonen wurde, auch wenn Nachfragen diesbezüglich kamen, nicht gesagt, wie sie sich in einer solchen Situation zu verhalten haben. Es wurde darauf hingewiesen, dass es den Versuchspersonen in jeder Situation freisteht, wie sie sich verhalten sollten.

Den Versuchspersonen der Unsicherheitsgruppe und den Versuchspersonen der „Keine Rückmeldung“ Gruppen wurde erklärt, dass die Automation nicht fehlerfrei ist und durchaus Automationsfehler machen kann. Dies wurde getan, da die Unsicherheitsgruppen durch die Einführung der Unsicherheitsrückmeldung implizit auf die Möglichkeit von Automationsfehlern schließen konnten und damit Sekundärvarianz durch Vorwissen erzeugt worden wäre. Alle Gruppen sollten somit mit dem gleichen Vorwissen bezüglich der Automation in den Versuch gehen.

Vor dem Beginn der ersten Versuchsfahrt wurden die Versuchspersonen instruiert, nach einem Startsignal die Automation zu aktivieren und das System auf die voreingestellte Geschwindigkeit beschleunigen zu lassen und erst dann mit der Bearbeitung der Nebenaufgaben zu beginnen. Sie wurden ebenso instruiert, dass sie so viele Nebenaufgaben wie möglich bearbeiten sollten, in dem Rahmen wie sie ein sicheres Fahren vertreten können. Dann wurde die erste Fahrt gestartet. Nach der ersten Fahrt bekamen die Versuchspersonen einen Fragebogen zum Vertrauen in die Automation zur Bearbeitung vorgelegt.

Danach wurde direkt die zweite Fahrt begonnen, für die die Instruktionen der ersten Fahrt noch einmal wiederholt wurden. Auch während der zweiten Fahrt wurden die Versuchspersonen in den Situationen mit einem Fremdfahrzeug nach ihrem situativen Vertrauen befragt. Nach Beendigung der zweiten Fahrt sollten die Versuchspersonen wieder den Fragebogen zum Vertrauen in die Automation sowie

einen Fragebogen zur subjektiv empfundenen Kooperation und einige weitere Fragen zur Akzeptanz der Automation bearbeiten.

Danach wurden die Versuchspersonen über den eigentlichen Zweck der Studie aufgeklärt, sie erhielten eine Aufwandsentschädigung und wurden verabschiedet.

## 7.3 Ergebnisse Studie 1

Da für die unterschiedlichen Hypothesen bzw. Experimentalpläne häufig die gleichen abhängigen Variablen genutzt werden, werden die Ergebnisse nicht je Hypothese bzw. Telexperiment, sondern je abhängiger Variable dargestellt.

### 7.3.1 Stichprobe

Insgesamt nahmen 28 Versuchspersonen an der ersten Studie teil. Die Versuchspersonen wurden sämtlich aus dem Probandenpool des DLR Instituts für Verkehrssystemtechnik akquiriert und erhielten eine Aufwandsentschädigung von 8 Euro / Std. Die Versuchspersonen wurden nach dem im Abschnitt „Generelles Vorgehen: Stichproben“ beschriebenen Schema auf die Bedingungen verteilt. Die genaue Aufteilung der Subgruppen ist Tabelle 5 zu entnehmen.

**Tabelle 5: Statistik der 4 Versuchspersonengruppen aus Studie 1**

Gruppe	Männlich / weiblich	Alter Min / Max	Alter Mittelwert / Stdabw	Führerschein Min / Max	Führerschein Mittelwert / Stdabw
Unsicherheit Ohne Fehler – mit Fehler	5 / 2	20 / 40	28,14 / 7,24	2,5 / 22	9,78
Unsicherheit Mit Fehler – ohne Fehler	5 / 2	20 / 38	27,29 / 6,55	2 / 16	7,33 / 5,00
Keine Rückmeldung Ohne Fehler – mit Fehler	5 / 2	19 / 49	28,71 / 11,74	2 / 32	12,75/ 13,22
Keine Rückmeldung Mit Fehler – ohne Fehler	3 / 4	20 / 61	29,00 / 14,72	4 / 43	12,28 / 14,17

### 7.3.2 Ergebnisse Kontrollierbarkeit

Die zuerst dargestellten Ergebnisse beziehen sich auf Effekte der Unsicherheitsrückmeldung auf die Kontrollierbarkeit eines Automationsfehlers. Die in

		Erste Fahrt																				Zweite Fahrt																							
Fahrt ohne Automationsfehler zuerst	Unsicherheitsrückmeldung	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
	Keine Rückmeldung	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Fahrt mit Automationsfehler zuerst	Unsicherheitsrückmeldung	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
	Keine Rückmeldung	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

### 7.3.2.1 Reaktion auf kritische Automationsfehler – Reaktionshäufigkeit

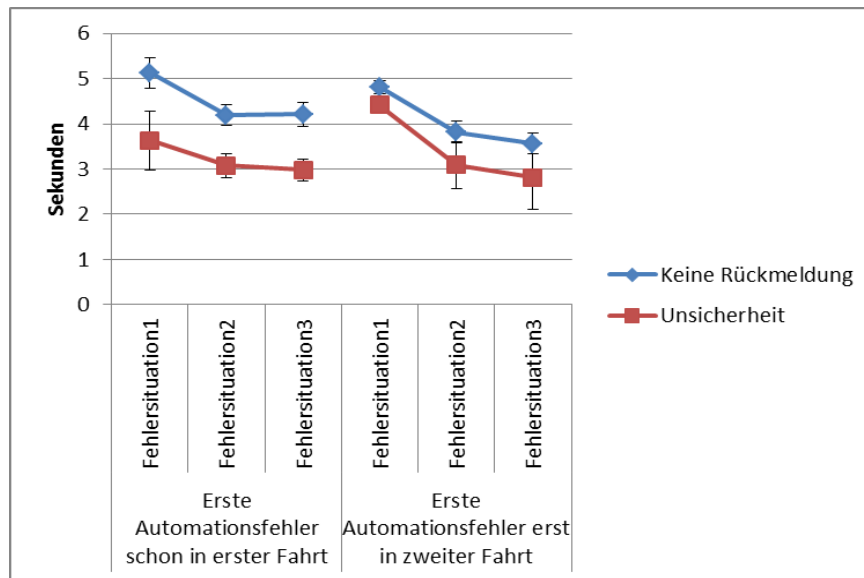
**Tabelle 6: Anteil der Versuchspersonen ohne Bremsreaktion in Automationsfehlersituation in Abhängigkeit von Rückmeldung und Fehlerzeitpunkt. Hinter dem Schrägstrich ist die jeweilige Gesamtanzahl der Versuchspersonen in dieser Bedingung aufgelistet.**

Aufgrund der geringen Fallzahlen (häufig 0) kann kein Chi Quadrattest durchgeführt werden, ein Test auf statistisch bedeutsame Unterschiede ist also nicht möglich. Augenscheinlich scheint jedoch die Unsicherheitsrückmeldung zu tendenziell häufigerem Eingreifen des Fahrers zu führen. Werden Automationsfehler in der ersten Fahrt erlebt, so zeigen alle Fahrer der Unsicherheitsbedingung bei allen drei Automationsfehlern eine erfolgreiche Kontrollreaktion. In der Bedingung keine Rückmeldung zeigt in jeder Fehlersituation jeweils eine Versuchsperson keine Reaktion.

91

### 7.3.2.2 Reaktion auf kritische Automationsfehler – Bremsreaktionszeit

Als maßgebliches Kriterium für die Güte der Kontrollreaktion wird die Bremsreaktionszeit gewählt (siehe Abschnitt „Messung des Cry Wolf Effektes“).



**Abbildung 28: Reaktionszeiten in den Situationen mit Automationsfehler (reagiert nicht auf Vorderfahrzeug). Aufgeteilt nach Faktor Rückmeldung (Keine Rückmeldung und Unsicherheit) und Faktor Fehlerzeitpunkt (erster Fehler in erster Fahrt & erster Fehler in zweiter Fahrt)**

### 7.3.2.3 Haupteffekte des Faktors Rückmeldung

Die Reaktionszeiten in der Unsicherheitsbedingung liegen, teilweise deutlich, unter denen der Bedingung keine Rückmeldung (Abbildung 28). Die Mixed Models Varianzanalyse zeigt einen deutlichen Haupteffekt des Faktors Rückmeldung mit einer mittleren Effektstärke (Tabelle 7). Der Unterschied in der Reaktionszeit zwischen der Unsicherheitsbedingung und keiner Rückmeldung scheint insbesondere dann besonders ausgeprägt zu sein, wenn weder die Erfahrung eines Fehlers vorliegt und der Zeitpunkt der Fehlererfahrung früh ist (erster Fehler in erster Fahrt).

**Tabelle 7: Ergebnisse Mixed Models ANOVA der Reaktionszeiten**

Faktor	F	df	Sig.	Eta
<b>Rückmeldung (Zwischensubjektfaktor)</b>	9.199	1 / 22	<b>.006</b>	,295
<b>Fehlererfahrung (Innersubjektfaktor)</b>	11.423	2 / 44	<b>.000</b>	,342
<b>Fehlerzeitpunkt</b>	.138	1 / 22	.714	
<b>Rückmeldung*Fehlererfahrung</b>	.011	2 / 44	.989	
<b>Rückmeldung*Fehlerzeitpunkt</b>	1.085	1 / 22	.309	
<b>Fehlererfahrung*Fehlerzeitpunkt</b>	.836	2 / 44	.440	
<b>Rückmeldung*Fehlererfahrung* Fehlerzeitpunkt</b>	.294	1 / 44	.745	

#### **7.3.2.4 Haupteffekte des Faktors Fehlererfahrung & Interaktion mit Rückmeldung**

Innerhalb einer Fahrt mit Fehlern zeigt sich ein deutlicher Effekt auf dem Messwiederholungsfaktor „Fehlererfahrung“, mit mittlerer Effektstärke (Tabelle 7). Wird ein Fehler ohne Vorerfahrung erlebt (erster Fehler in einer Fahrt) so ist die Reaktionszeit insgesamt höher, verglichen mit der Reaktionszeit ohne Fehlererfahrung.

Es können auch hier keine Interaktionseffekte zwischen dem Faktor Fehlererfahrung und dem Faktor Rückmeldung beobachtet werden. In beiden Bedingungen sinkt die Reaktionszeit vom ersten zum zweiten erlebten Fehler um etwa den gleichen Wert ab. Dies gilt für die Fehler in der ersten Fahrt als auch für die Fehler in der zweiten Fahrt. Der Reaktionszeitunterschied zwischen Unsicherheit und keiner Rückmeldung bleibt in etwa gleich.

Des Weiteren lassen sich auch keine weiteren Interaktionseffekte, weder eine Interaktion von Fehlererfahrung x Fehlerzeitpunkt noch von Fehlersituation x Gruppe x Reihenfolge (Tabelle 7) erkennen.

#### **7.3.2.5 Haupteffekt des Faktors Fehlerzeitpunkt & Interaktion mit Rückmeldung**

Diese Überprüfung bezieht sich auf die Anfälligkeit der Unsicherheitsrückmeldung für einen Cry Wolf Effekt. Insgesamt zeigt sich kein Effekt der Vorerfahrung einer fehlerfreien Automation auf die Reaktionszeit (Faktor Fehlerzeitpunkt) und auch kein Interaktionseffekt der Faktoren Rückmeldung und Fehlerzeitpunkt (Tabelle 7). Dies scheint zunächst gegen einen Cry Wolf Effekt zu sprechen. In der Unsicherheitsbedingung zeigt sich jedoch augenscheinlich ein deutlicher Mittelwertunterschied der Reaktionszeit zwischen dem ersten Fehler der in der ersten Fahrt erlebt wird und dem ersten Fehler der erst in der zweiten Fahrt erlebt wird. Wird

zunächst über eine längere Periode trotz mehrfacher Unsicherheitsrückmeldung kein Fehler erlebt, so fällt die Reaktionszeit im Schnitt deutlich höher aus, verglichen mit der Reaktion bei früh in der Fahrt erlebtem Fehler.

Um die Anfälligkeit der Unsicherheitsrückmeldung für einen Cry Wolf Effekt genauer betrachten zu können, wird eine separate Analyse der Effekte von Rückmeldungsart und Fehlererfahrung nur für die ersten in einer Fahrt erlebten Fehler durchgeführt.

**Tabelle 8: Mittelwerte und Standardfehler der Reaktionszeiten bei erstem erlebtem Automationsfehler in Abhängigkeit von Rückmeldung und Fehlererfahrungszeitpunkt.**

	Erster Fehler in erster Fahrt	Erster Fehler in zweiter Fahrt
Unsicherheit	MW = 3,63157 STDERR = 0,69678	MW = 4,4136 STDERR = 0,19786
Keine Rückmeldung	MW = 5,1235 STDERR = 0,36745	MW = 4,8101 STDERR = 0,16199

Die zweifaktorielle ANOVA mit den Faktoren Rückmeldung und Fehlerzeitpunkt ergibt einen signifikanten Haupteffekt der Rückmeldung, jedoch anders als Abbildung 28 vermuten lässt, keinen Effekt des Fehlerzeitpunktes und auch keinen Interaktionseffekt aus Fehlerzeitpunkt und Rückmeldung (Tabelle 9).

**Tabelle 9: Ergebnisse der ANOVA für die Reaktionszeit in der jeweils ersten Situation mit Automationsfehler.**

	df	F	Sig.	Partielles Eta
Rückmeldung	1, 21	4.412	.048	.174
Fehlerzeitpunkt	1, 21	.272	.608	.013
Rückmeldung x Fehlerzeitpunkt	1, 21	1.485	.237	.066

Des weiteren werden paarweise Vergleiche der Reaktionszeiten durchgeführt:

- Einmal jeweils auf Ebene der Unsicherheitsrückmeldung und der Gruppe keine Rückmeldung, zwischen den beiden Fehlerzeitpunkten.
- Einmal jeweils auf Ebene des frühen Fehlerzeitpunktes und des späten Fehlerzeitpunktes, zwischen den beiden Rückmeldungsbedingungen.

Aufgrund der kleinen Stichprobengröße (N= 6 – 7) werden nonparametrische Mann-Whitney U Tests durchgeführt (Tabelle 10).

Beim Vergleich erster Fehler in erster Fahrt mit erstem Fehler in zweiter Fahrt innerhalb der Unsicherheitsgruppe, kann kein signifikanter Unterschied zwischen der Reaktionszeit festgestellt werden. Der vermeintliche Unterschied auf deskriptiver

Ebene muss demnach zunächst als rein zufällig angesehen werden. Ein Cry Wolf Effekt lässt sich so nicht zweifelsfrei identifizieren. Den Tests zufolge ist der Unterschied zwischen Unsicherheit und keine Rückmeldung bei dem ersten Fehler in der ersten Fahrt jedoch nur knapp nicht signifikant. Die Tests zeigen aber auch, dass es zwischen den Reaktionszeiten der Bedingungen Unsicherheit und keine Rückmeldung bei dem ersten Fehler in der zweiten Fahrt nichtmals einen tendenziell bedeutsamen Unterschied gibt.

**Tabelle 10: Ergebnisse der Mann-Whitney U Test der Reaktionszeitvergleiche bei erstem erlebtem Automationsfehler**

Vergleich	p
Unsicherheit – Fehler in erster Fahrt vs. Unsicherheit Fehler in zweiter Fahrt	.445
Keine Rückmeldung – Fehler in erster Fahrt vs. kein Rückmeldung Fehler in zweiter Fahrt	.589
Unsicherheit vs Keine Rückmeldung bei Fehler in erster Fahrt	.073
Unsicherheit vs. keine Rückmeldung bei Fehler in zweiter Fahrt	.240

### 7.3.3 Ergebnisse Gesamtvertrauen:

#### 7.3.3.1 Auswirkung auf die Entstehung adäquaten Vertrauens

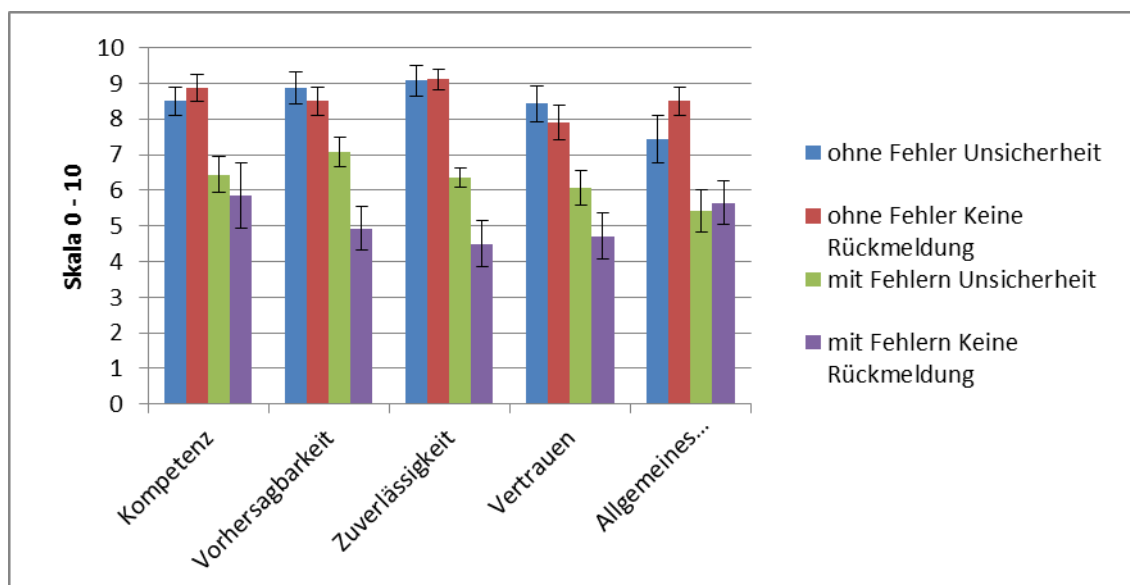
Im ersten Schritt wird die Auswirkung der Unsicherheitsrückmeldung auf die Bildung eines adäquaten Vertrauens nach einer Fahrt ohne Automationsfehler betrachtet (Abbildung 29 grün markiert). Dazu wird ein Vergleich mit der Gruppe keine Rückmeldung angestellt.

		Erste Fahrt		Zweite Fahrt	
Fahrt ohne Automationsfehler zuerst	Unsicherheitsrückmeldung	Fahrt ohne Fehler	Fragebogen Gesamttrust	Fahrt mit Fehler	Fragebogen Gesamttrust
	Keine Rückmeldung	Fahrt ohne Fehler	Fragebogen Gesamttrust	Fahrt mit Fehler	Fragebogen Gesamttrust
Fahrt mit Automationsfehler zuerst	Unsicherheitsrückmeldung	Fahrt mit Fehler	Fragebogen Gesamttrust	Fahrt ohne Fehler	Fragebogen Gesamttrust
	Keine Rückmeldung	Fahrt mit Fehler	Fragebogen Gesamttrust	Fahrt ohne Fehler	Fragebogen Gesamttrust

**Abbildung 29: Vergleich des Gesamttrust nach der ersten Fahrt für die Bedingung mit Fehler und ohne Fehler jeweils für beide Rückmeldungsarten.**

Deskriptiv zeigt sich, dass die Trustratings in beiden Rückmeldungsgruppen auf allen fünf Dimensionen recht hoch liegen (Min 7,4286 und Max 9,1143 auf einer Skala von 0-10 – siehe Abbildung 30).

Eine multivariate Varianzanalyse für die Faktorstufen Unsicherheit und keine Rückmeldung auf den fünf Trust-Dimensionen, zeigt keinen signifikanten Unterschied in den Trustratings zwischen der Unsicherheitsgruppe und der „Keine Rückmeldung“ Gruppe (Tabelle 11). Lediglich in der Dimension „Allgemeines Vertrauen“ gibt es einen schwachen Effekt (.140 partielles Eta), der sich etwas deutlicher von den anderen Dimensionen abhebt, jedoch ebenfalls nicht signifikant ist (Tabelle 11).



**Abbildung 30: Durchschnittliches subjektives Gesamtvertrauen in Unsicherheits- und „Keine-Rückmeldung“ Gruppe nach Fahrt ohne und Fahrt mit Fehlern**

**Tabelle 11: Test der Mittelwertunterschiede in den einzelnen Trustdimensionen nach Fahrt ohne Automationsfehler.**

Trustdimension	df	F	Sig.	Partielles Eta
Kompetenz	1, 12	.434	.523	.035
Vorhersagbarkeit	1, 12	.361	.559	.029
Zuverlässigkeit	1, 12	.007	.937	.001
Vertrauen	1, 12	.574	.463	.046
Vertrauen allgemein	1, 12	1.951	.188	.140

### 7.3.3.2 Auswirkungen auf die Entstehung von Misstrust

Im nächsten Schritt wird betrachtet, in wie weit sich eine Unsicherheitsrückmeldung gegen ein übermäßiges Absinken des Vertrauens nach Fehlererlebnissen im Sinne eines Misstrust auswirkt (Hypothese 4b). Dazu wird das nach der Fahrt mit



Automationsfehlern erhobene Vertrauen zwischen der Unsicherheitsbedingung und der Bedingung „keine Rückmeldung“ verglichen (Abbildung 29 rot markiert).

Auf der deskriptiven Ebene zeigen sich in einigen Trustdimensionen bereits deutliche Mittelwertunterschiede (Abbildung 30). Höhere Trustratings sind insbesondere in den Trustdimensionen Vorhersagbarkeit, Zuverlässigkeit und Vertrauen zu beobachten. So schätzen die Versuchspersonen im Durchschnitt eher ein, dass die Automation mit Unsicherheitsrückmeldung das Verhalten mit einem Mindestgrad an Sicherheit vorhersagen kann (Vorhersagbarkeit). Außerdem wird von den Versuchspersonen die Automation eher als fehlerfrei beurteilt, wenn Unsicherheit zurückgemeldet wurde (Zuverlässigkeit) und dies obwohl die gleiche Anzahl an Automationsfehlern erlebt wurde, wie in der Vergleichsgruppe „keine Rückmeldung“.

Dies könnte ein Hinweis dahingehend sein, dass Unsicherheitsrückmeldung vor unangepasstem Automationsverhalten dazu führt, dass dieses Verhalten nicht so streng als Fehler beurteilt wird. Die Versuchspersonen haben auch tendenziell eher Vertrauen darin, dass die Automation bestimmte Aufgaben effektiv bearbeitet, ohne dass sie das prüfen können (Vertrauen).

Eine multivariate Varianzanalyse ergibt bei den genannten Dimensionen auch deutliche (Vorhersagbarkeit und Zuverlässigkeit) oder zumindest tendenzielle Unterschiede (Vertrauen) zwischen der Unsicherheitsgruppe und der „Keine Rückmeldung“ Gruppe (Tabelle 12).

In den Dimensionen Kompetenz (in welchem Ausmaß erfüllt die Automation die Fahraufgabe effektiv?) und dem allgemeinen Vertrauen (in welchem Ausmaß vertrauen sie der Automation im Allgemeinen) gibt es nicht einmal tendenzielle Unterschiede zwischen den Rückmeldungsgruppen (Tabelle 12).

**Tabelle 12: Test der Mittelwertunterschiede in den einzelnen Trustdimensionen nach Fahrt mit mehreren Automationsfehlern.**

Trustdimension	df	F	Sig.	Partielles Eta
Kompetenz	1, 12	.301	.593	.024
Vorhersagbarkeit	1, 12	8.411	<b>.013</b>	.412
Zuverlässigkeit	1, 12	7.116	<b>.021</b>	.372
Vertrauen	1, 12	2.793	.120	.189
Vertrauen allgemein	1, 12	.063	.806	.005

### 7.3.3.3 Auswirkung der Erfahrung von Automationsfehlern auf das Vertrauen

Zuletzt werden beide Fahrten, die mit Fehler und die ohne Fehler, miteinander verglichen. So soll die Auswirkung der Erfahrung echter Fehler auf das Gesamtvertrauen innerhalb der Unsicherheitsbedingung überprüft werden. Es wird also auf einen Haupteffekt des Faktors „Fehlerfreiheit“ innerhalb der Bedingung Unsicherheit getestet.

Werden die Vertrauensratings innerhalb der Unsicherheitsgruppe betrachtet, so fällt auf, dass diese von der Bedingung keine Fehler zur Bedingung Fehler deutlich sinken.

Die MANOVA des Faktors „Fehlerfreiheit“ auf der Ebene der Unsicherheitsrückmeldung kann einen deutlichen Haupteffekt auf allen 5 Trustdimensionen aufzeigen (Tabelle 13). Werden Fehler der Automation erlebt, so sinkt trotz der Rückmeldung von Automationsunsicherheit, das Vertrauen in die Automation.

**Tabelle 13: Ergebnisse der MANOVA Test auf Haupteffekt des Faktors „Fehlerfreiheit“ innerhalb der Unsicherheitsbedingung.**

Trustdimension	Df	F	Sig.	Partielles Eta
Kompetenz	1, 12	10.469	.007	.466
Vorhersagbarkeit	1, 12	8.333	.014	.410
Zuverlässigkeit	1, 12	27.948	.000	.700
Vertrauen	1, 12	11.423	.005	.488
Vertrauen allgemein	1, 12	5.102	.043	.298

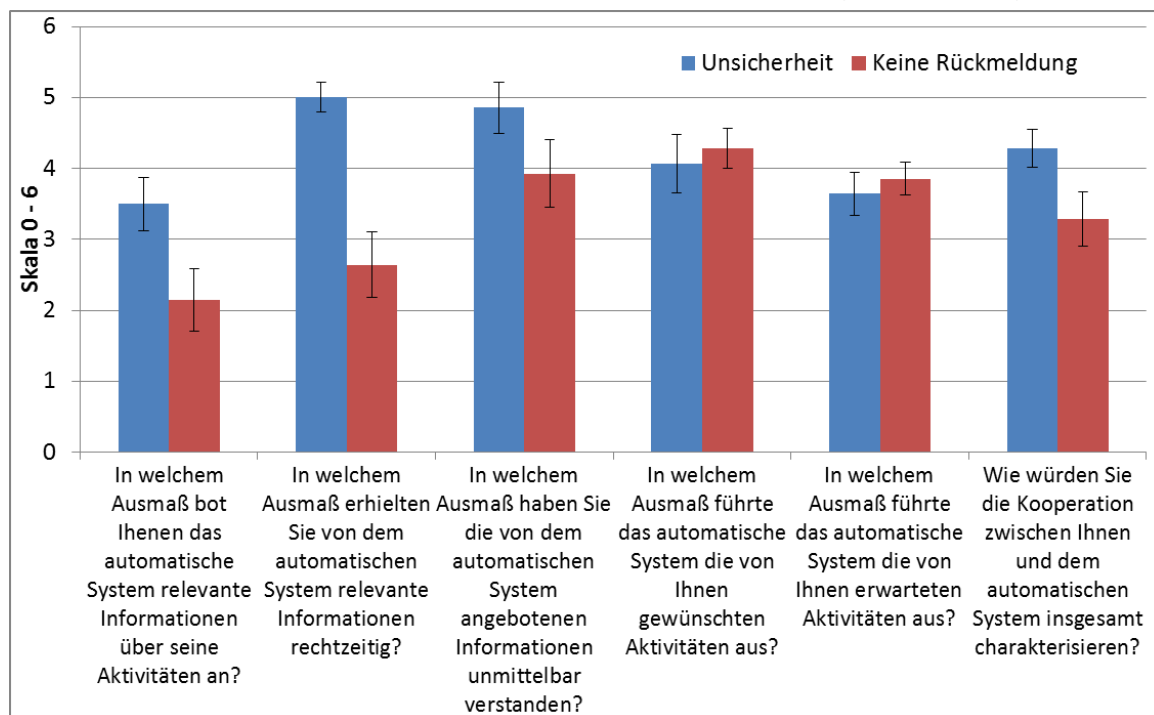
### 7.3.4 Ergebnisse: Akzeptanz

Zusätzlich zum subjektiv empfundenen Vertrauen in die Automation werden als weitere Akzeptanzmaße die subjektiv empfundene Kooperation zwischen Fahrer und Automation und die Tendenz, die Automation nach den gemachten Erfahrungen auszuschalten, bzw. weiter zu nutzen, erhoben.

#### 7.3.4.1 Subjektiv empfundene Kooperation

Nach Abschluss der zweiten Fahrt wurde ein Fragebogen vorgelegt, der verschiedene Aspekte der Kooperation zwischen Fahrzeugautomation und Fahrer erfassen sollte (siehe Erfassung der Akzeptanz in die Automation). Die Ergebnisse der Befragung zeigen, dass insbesondere bezüglich der von der Automation erhaltenen Informationen, die Automation mit Unsicherheitsrückmeldung als deutlich

besser beurteilt wird. Auch in der Beurteilung der Kooperation zwischen Fahrer und Fahrzeugautomation insgesamt wird die Automation mit Unsicherheitsrückmeldung besser beurteilt als die Kontrollgruppe „Keine Rückmeldung“ (Abbildung 31).



**Abbildung 31: Subjektiv empfundene Kooperation zwischen Fahrer und Automation nach der zweiten Fahrt (0 = gar nicht/sehr schlecht - 6 = immer/sehr gut)**

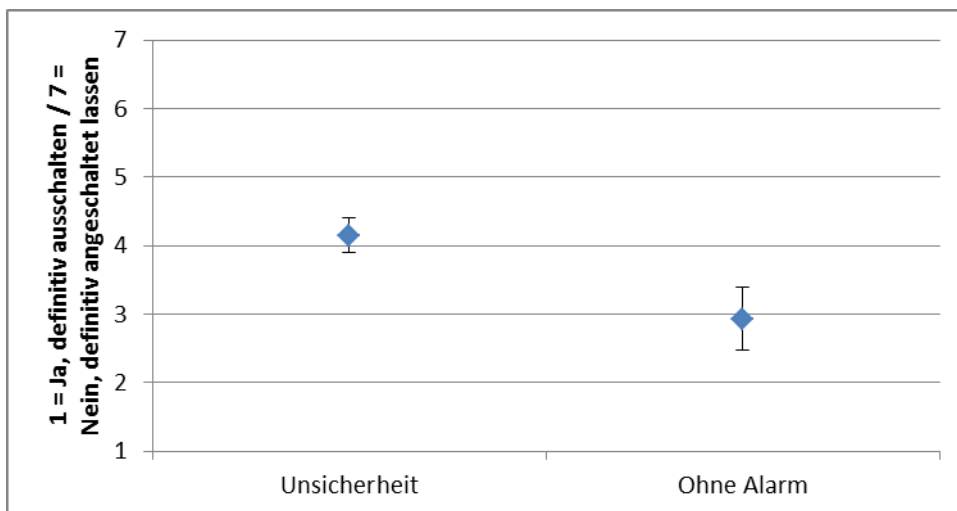
Mittels einer MANOVA wird auf Unterschiede zwischen Unsicherheitsrückmeldung und „Keine Rückmeldung“ in den einzelnen Kooperationsfragen getestet. Demnach beurteilten die Versuchspersonen, dass in der Automation mit Unsicherheitsrückmeldung diese in einem größeren Ausmaß Informationen über dessen Aktivitäten anbot und, dass diese Informationen in einem deutlich stärkeren Ausmaß rechtzeitig kamen. Insgesamt wurde die Kooperation zwischen Fahrer und Fahrzeug mit Unsicherheitsrückmeldung als deutlich besser beurteilt.

**Tabelle 14: Ergebnisse der MANOVA der einzelnen Kooperationsdimensionen.**

Frage Nr.	df	F	Sig.	Partielles Eta
Koop 1	1 / 26	5,476	<b>,027</b>	,174
Koop 2	1 / 26	21,418	<b>,000</b>	,452
Koop 3	1 / 26	2,428	,131	,085
Koop 4	1 / 26	,183	,673	,007
Koop 5	1 / 26	,310	,582	,012
Koop 6	1 / 26	4,583	<b>,042</b>	,150

### 7.3.4.2 Nutzungstendenz – Einschalten vs. Ausschalten

Ein wichtiger Aspekt der Akzeptanz ist die Auswirkung der Erfahrung mit einer Fahrzeugautomation auf die Tendenz, die Fahrzeugautomation weiter nutzen zu wollen. In jeder Versuchspersonengruppe wurden die Versuchspersonen nach Beendigung der zweiten Fahrt gefragt, ob sie aufgrund der gemachten Erfahrung das System eher eingeschaltet lassen würden oder eher ausschalten würden. Abbildung 32 zeigt, dass die Versuchspersonen der Unsicherheitsgruppe die Automation eher eingeschaltet lassen würden als die Versuchspersonen der Kontrollgruppe „Keine Rückmeldung“.



**Abbildung 32: Subjektive Tendenz die Fahrzeugautomation aufgrund der gemachten Erfahrung eingeschaltet zu lassen oder auszuschalten.**

Zum Test des Unterschieds zwischen Unsicherheits- und „Keine Rückmeldung“ wird ein  $t$ -Test für unabhängige Stichproben durchgeführt. Dieser ergibt einen signifikanten Unterschied  $t(26) = 2,302$ ,  $p .030$ .

## 7.4 Diskussion Studie 1

Studie eins sollte einen Überblick über Effekte einer Rückmeldung des Automationszustandes in Form von Automationsunsicherheit liefern. Dazu sollte untersucht werden, wie sich das Vorhandensein einer Unsicherheitsrückmeldung, die Erfahrung eines Automationsfehlers, die erfahrene Automationsreliabilität (Zeit bzw. Anzahl der Situationen ohne Fehler) und die Menge an erlebten „falschen Alarmen“, auf die Erwartung eines Automationsfehlers auswirken.

Wesentliche Bestandteile waren:

1. Die Überprüfung, ob eine Unsicherheitsrückmeldung überhaupt effektiv ist und einen Vorteil gegenüber gar keiner Rückmeldung hat (Fehlererwartung durch Rückmeldung),

2. Die Untersuchung, ob eine Unsicherheitsrückmeldung generell anfällig gegenüber einem Cry Wolf Effekt ist (Einfluss der erlebten „falschen Alarme“ auf die Effizienz der Unsicherheitsrückmeldung).
3. Die moderierenden Effekte einer Unsicherheitsrückmeldung auf die Bildung des Vertrauens in die Automation.
4. Die Auswirkungen auf den Grad der wahrgenommenen Kooperation.

Zu diesen unterschiedlichen Untersuchungsschwerpunkten wurden in Abschnitt „Ziele Studie 1“ Untersuchungsziele und zugehörige Forschungsfragen und Hypothesen formuliert. Der Diskussionsteil wird daher nach den einzelnen Teilzielen der Studie 1 gegliedert.

**Ziel 1:** Überprüfung der Effektivität einer Unsicherheitsrückmeldung – hat eine Unsicherheitsrückmeldung überhaupt einen Vorteil gegenüber keiner Rückmeldung?

Bei den betrachteten Automationsfehlersituationen kann ein Haupteffekt des Faktors Rückmeldung auf die Reaktionszeit beobachtet werden. In Abhängigkeit der weiteren Faktoren Fehlerzeitpunkt (bzw. Anzahl erlebter „falscher Alarme“) und Fehlererfahrung, können in der Unsicherheitsbedingung teils deutlich bessere Reaktionszeiten erzielt werden als in den Vergleichsbedingungen ohne Rückmeldung.

Eine Unsicherheitsmeldung ist demnach, trotz der vermittelten Unsicherheit bezüglich des Eintretens eines Automationsfehlers, geeignet eine situative Fehlererwartung zu erzeugen, die zu einer besseren Kontrollierbarkeit führt. Sie ist somit als effektiv zu bezeichnen und das notwendige Kriterium zur weiteren Untersuchung der Unsicherheitsrückmeldung ist damit erfüllt.

**Ziel 2:** Untersuchung der Cry Wolf Anfälligkeit einer Unsicherheitsrückmeldung in Abhängigkeit der Häufigkeit der erlebten falschen Alarme sowie Untersuchung der Auswirkungen des Erlebens einer fehlerfreien Automation auf Overtrust und Kontrollierbarkeit eines am Ende stattfindenden Automationsfehlers.

Dazu wird die Kontrollierbarkeit des ersten erlebten Automationsfehlers betrachtet. Dieser wird, je nach Bedingung, zu Anfang der ersten Fahrt erlebt (früher Fehlerzeitpunkt), oder erst, nachdem im ersten Fahrtabschnitt eine fehlerfreie Automation und eine Reihe von falschen Alarmen erlebt wurde (später Fehlerzeitpunkt).

Dabei zeigt sich, dass in der Unsicherheitsbedingung die Kontrollierbarkeit im früh erlebten Automationsfehler und einer entsprechend geringen Anzahl vorher erlebter

falscher Alarme, besser zu sein scheint, als bei einem Automationsfehler der spät erlebt wird und vorher eine größere Menge falscher Alarme erlebt wurden.

Die Reaktionszeit in der Bedingung keine Rückmeldung ist zu beiden Fehlerzeitpunkten in etwa gleich. Da in dieser Bedingung keine Rückmeldungen erlebt werden, welche auf potenzielle Automationsfehler hinweisen könnten, wird bis zum ersten Automationsfehler eine scheinbar fehlerfreie Automation erlebt.

Da die Reaktionen der Fahrer unabhängig davon sind, wie lange die scheinbar fehlerfreie Automation erlebt wurde, scheint es so, als hätte das Erleben einer fehlerfreien Automation keine Auswirkungen auf die Erwartung eines Automationsfehlers. In diesem Kontext ist es aber wahrscheinlicher, dass bereits bei frühem Erleben eines Automationsfehlers ein Automationsfehler gar nicht erwartet wird, wenn die Erwartung nicht situativ durch eine Rückmeldung erhöht wird.

Trotz der augenscheinlichen Wechselwirkung von Fehlerzeitpunkt und Rückmeldung auf die Reaktionszeit, kann kein signifikanter Interaktionseffekt beobachtet werden. Dies ist mit hoher Wahrscheinlichkeit der geringen Stichprobengröße geschuldet, die einen wohlmöglich vorhandenen Effekt nicht identifizierbar macht.

Auch die durchgeführten non-parametrischen Tests, können in der Unsicherheitsbedingung keinen Unterschied in der Reaktionszeit zwischen frühem Fehlerzeitpunkt und spätem Fehlerzeitpunkt herausstellen.

Obwohl sich der Cry Wolf Effekt nicht statistisch absichern ließ, lässt die Systematik der Ergebnisse den Schluss zu, dass, anders als in der Hypothese behauptet, die Unsicherheitsrückmeldung nicht immun gegen einen Cry Wolf Effekt ist.

In Studie 1 wurden allerdings auch extreme Bedingungen umgesetzt, da, trotz des recht kurzen Zeitraums in dem die Simulatorfahrt durchgeführt wurde, ein bestimmtes Potenzial für einen Cry Wolf Effekt erzeugt werden musste. Die Menge an erlebten falschen Alarmen bis zum ersten Automationsfehler und auch das Verhältnis aus klaren Situationen ganz ohne Rückmeldung zu unklaren Situationen mit Unsicherheitsrückmeldung von fast 1:1 ist insgesamt unnatürlich hoch. Wenn jedoch überhaupt keine Cry Wolf Anfälligkeit der Unsicherheitsrückmeldung vorliegen würde, hätte auch die geschilderte übertriebene Umsetzung keinen negativen Effekt gehabt. In Folgestudien soll daher eine Betrachtung der Cry Wolf Anfälligkeit relativ zu anderen Rückmeldestrategien stattfinden.

So ist es möglich, dass die Cry Wolf Anfälligkeit einer Unsicherheitsrückmeldung, im Vergleich zu alternativen Rückmeldungen, relativ gering ausfällt. Ebenso sollte eine Betrachtung unter realistischeren Häufigkeiten falscher Alarme stattfinden. Damit können Aussagen darüber getroffen werden, ob die verbleibende Anfälligkeit der Unsicherheitsrückmeldung für den Cry Wolf Effekt einen Einsatz im Fahrzeug trotzdem rechtfertigt.

**Ziel 3:** Untersuchung der Auswirkungen einer Fehlererfahrung auf die generelle Erwartung eines Fehlers und die Effektivität einer Unsicherheitsrückmeldung.

Die Erfahrung eines Automationsfehlers scheint insgesamt einen fast ebenso deutlichen Einfluss auf die Erwartung von Automationsfehlern zu haben, wie die Unsicherheitsrückmeldung. So wird die Kontrollierbarkeitsleistung auch ohne Rückmeldung insgesamt deutlich verbessert, wenn vorher bereits ein Automationsfehler erlebt wurde. Wird zusätzlich zu der Erfahrung eines Automationsfehlers bei nachfolgenden Automationsfehlern Unsicherheit zurückgemeldet, so ist die Kontrollierbarkeit noch besser.

Die Erfahrung durch das Erleben von Automationsfehlern bildet anscheinend geringeres Vertrauen in die Automation und somit eine höhere Grunderwartung von Automationsfehlern, welche die Effektivität einer Unsicherheitsrückmeldung unterstützt. Darauf weist der parallele Verlauf der Reaktionszeitentwicklung hin. Die Reaktionszeit der Bedingung Unsicherheit liegt immer im etwa gleichen Maß unter der Reaktionszeit der Bedingung keine Rückmeldung. Allgemeine Fehlererwartung durch Fehlererfahrung und situative Fehlererfahrung durch Rückmeldung scheinen demnach, wie es auch in der Hypothese behauptet wird, additiv zu wirken.

Bezüglich des beobachteten möglichen Cry Wolf Effektes der Unsicherheitsbedingung scheint dieser die Erwartung eines Automationsfehlers nicht vollständig zu unterdrücken.

Auch unter Wirkung des Cry Wolf Effektes erzeugt die Unsicherheitsrückmeldung, gemessen an der Reaktion auf den Fehler, anscheinend genau so viel Erwartung eines Automationsfehlers, wie sie durch vorherige Erfahrung von Automationsfehlern in der Bedingung ohne Rückmeldung resultieren. Dies gilt nur innerhalb des im Rahmen dieser Studie erzeugten Ausmaßes des Cry Wolf Effektes.

**Ziel 4:** Untersuchung der Auswirkungen einer Unsicherheitsrückmeldung auf das Vertrauen in die Automation im Vergleich zu keiner Rückmeldung.

Wenn das Vertrauen nach einer fehlerfreien Fahrt erhoben wird, kann, zumindest mittels des eingesetzten Fragebogens, keine Veränderung des Gesamtvertrauens durch eine Unsicherheitsrückmeldung allein beobachtet werden. Die Vertrauensbeurteilungen unterscheiden sich in keiner Vertrauensdimension signifikant von denen der Kontrollgruppe „Keine Rückmeldung“ und befinden sich durchweg auf einem recht hohen Niveau. Allenfalls in der Dimension „Allgemeines Vertrauen“, definiert durch die Frage „In welchem Ausmaß vertrauen Sie der Automation im Allgemeinen?“ ist ein tendenzieller Effekt beobachtbar. Die wiederholte Rückmeldung von Automationsunsicherheit ohne das tatsächliche

Erleben von Automationsfehlern führt also nicht, wie in der Hypothese behauptet, zu einem niedrigeren, aber dafür angepassten Vertrauen, verglichen mit einer Automation, in der keine Rückmeldung gegeben wird.

Es kann alternativ sein, dass gerade, weil nach der Unsicherheitsrückmeldung kein Fehler folgt, das Vertrauen in die Automation hoch ist, da trotz der Warnung die Automation keine Fehler macht. Hier muss wahrscheinlich deutlicher zwischen einem Vertrauen in das Rückmeldesystem und dem Vertrauen in die Automation insgesamt unterschieden werden.

Die Erfahrung von Fehlern hingegen, führt zu einem deutlichen Absinken des Gesamtvertrauens in die Automation, und zwar auf allen 5 erfassten Dimensionen. Werden Fehler erlebt, ist das Vertrauen mit Unsicherheitsrückmeldung zwar deutlich niedriger als nach einer Fahrt ohne Fehler, jedoch, über alle Trust-Dimensionen betrachtet, zumindest tendenziell besser als in der Bedingung ohne Rückmeldung. In den beiden Trustdimensionen Vorhersagbarkeit und Zuverlässigkeit ist das Vertrauen in eine Automation mit Unsicherheitsrückmeldung tatsächlich signifikant besser. Dabei wird die Vertrauensdimension „Vorhersagbarkeit“ durch die Frage definiert „In welchem Ausmaß können Sie das Verhalten der Automation mit einem Mindestgrad an Sicherheit vorhersagen?“

Die Vorhersagbarkeit des Verhaltens einer Automation scheint eine wesentliche Komponente des Vertrauens in die Automation zu sein. Eine Rückmeldung, welche die Erwartung von Automationsfehlern situativ erhöht, trägt entsprechend zu einer Steigerung der Vorhersagbarkeit von Automationsverhalten bei. Die Ergebnisse passen zu den Beobachtungen anderer Autoren, dass eine grundlegende Erwartung von Automationsfehlern, z.B. vermittelt durch Kenntnisse zur Reliabilität der Automation, das Vertrauen der Automation anpassen kann (z.B. Bagheri & Jamieson (2004). Lee & See (2004) behaupteten dazu passend, dass sich das Vertrauen in die Automation, gemäß den zur Verfügung stehenden Informationen über das Automationsverhalten und die Automationskompetenz, entwickelt.

Die Unsicherheitsrückmeldung wirkt sich, über die reine Vorhersagbarkeit hinaus, aber auch signifikant auf die Vertrauensdimension „Zuverlässigkeit“ aus, welche durch die Frage „In welchem Ausmaß ist die Automation fehlerfrei“ definiert ist. Versuchspersonen nehmen eine Automation, die Unsicherheit zurückmeldet, anscheinend als fehlerfreier wahr, verglichen mit einer Automation ohne Rückmeldung, und dies, obwohl Versuchspersonen die gleiche Art und Anzahl an Fehlern erlebt haben.

Eine mögliche Erklärung ist die, dass durch die Unsicherheitsrückmeldung eine Begründung für das Auftreten des Automationsfehlers vermittelt wird, nämlich, weil die Automation in dieser Situation unsicher ist. Eine solche Begründung für ein nicht



angepasstes Verhalten könnte dazu führen, dass das Verhalten gar nicht als Fehler angesehen wird, sondern tatsächlich als Bestandteil des normalen Verhaltens. Für eine verlässliche Aussage dazu muss jedoch ein Vergleich zwischen der Unsicherheitsrückmeldung und einer konventionellen Alarmgestaltung stattfinden, welche nur einen warnenden, jedoch keinen begründenden Charakter hat.

Zusammengefasst passt das Erleben von Automationsfehlern das Vertrauen deutlich besser an, als nur das Erleben von Unsicherheitsrückmeldung ohne Automationsfehler, wenn in einer Fahrt kein tatsächlicher Automationsfehler erlebt wird. Die Hypothese konnte demnach nicht bestätigt werden.

Werden jedoch Automationsfehler in einer Fahrt erlebt, wirkt sich die Unsicherheitsrückmeldung in einem gewissen Grad puffernd auf die Entstehung von Misstrust aus, was die zu diesem Aspekt formulierte Hypothese zumindest teilweise stützen kann.

In wie weit die beobachtete puffernde Wirkung aber tatsächlich eine Eigenschaft der Unsicherheitsrückmeldung ist und nicht auch durch konventionelle Arten der Rückmeldungsgestaltung hätte erreicht werden können, kann ohne einen Vergleich zu einer solchen alternativen Rückmeldung nicht geklärt werden. In wie weit die erlebten falschen Alarme mit unterschiedlichen Rückmeldungsarten auf die Entstehung des Vertrauens wechselwirken, ist Gegenstand der Studien 2 & 3.

**Ziel 5:** Untersuchung der Auswirkung der Unsicherheitsrückmeldung auf die subjektiv wahrgenommene Kooperation mit der Automation und die Akzeptanz des Gesamtsystems.

Eine Fahrzeugautomation die Fehler macht, wird insgesamt als kooperativer wahrgenommen, wenn sie vor diesen Fehlern dem Fahrer eine Information bezüglich der Fehlermöglichkeit zurückmeldet. In den erhobenen Fragen zur Kooperation äußert sich dies nachvollziehbarerweise darin, dass die Versuchspersonen eher davon überzeugt sind, dass die Automation mit Unsicherheitsrückmeldung relevante Informationen zurückmeldet und dies auch rechtzeitig tut, als dies für die Bedingung ohne Rückmeldung der Fall ist. Dies widerspiegelt auch die Ergebnisse zur Vertrauensdimension Vorhersagbarkeit.

Direkt nach dem Grad der Kooperativität befragt, beurteilen die Versuchspersonen die Automation mit Unsicherheitsrückmeldung auch als deutlich kooperativer als die Automation ohne Rückmeldung. Obwohl die Unsicherheitsrückmeldung einen sehr unverbindlichen Charakter hat und nur nach einem Teil der erfolgten Unsicherheitsrückmeldungen tatsächlich ein Automationsfehler folgt, wird die subjektiv empfundene Kooperation als sehr hoch eingeschätzt.

Auch bezüglich der Tendenz, die Automation weiter zu nutzen oder auszuschalten, lassen sich die Ergebnisse, mit denen zur subjektiv empfundenen Kooperation in Einklang bringen. Insgesamt würden die Versuchspersonen die Automation mit Unsicherheitsrückmeldung eher eingeschaltet lassen als eine Automation ohne Rückmeldung. Die Nutzungstendenz ist zwar auch in der Unsicherheitsgruppe insgesamt eher schlecht, dies ist aber wohl auf die laborbedingte extrem hohe Anzahl an erlebten Warnungen und Fehler in recht kurzem Zeitraum zurückzuführen. Eine Automation die Fehler macht, scheint demnach deutlich besser akzeptiert zu werden, wenn vor diesen Fehlern eine Rückmeldung bezüglich der Fehlermöglichkeit kommt. Dies scheint auch zu gelten, wenn die Rückmeldung einen unverbindlichen Charakter hat und insgesamt deutlich häufiger ohne einen nachfolgenden Automationsfehler erlebt wurde.

**Zusammengefasst** lassen die Ergebnisse der ersten Studie den Schluss zu, dass eine Unsicherheitsrückmeldung insgesamt deutlich positive Effekte auf die Kontrollierbarkeit von Automationsfehlern, die Entstehung von Misstrust und die Akzeptanz von Automationsfehlern, haben kann. Damit wird eines der primären Ziele von Studie eins erreicht, nämlich zu klären, ob die Unsicherheitsrückmeldung als unverbindliche Rückmeldung überhaupt zu einer Verbesserung der Kontrollierbarkeit führen kann. Dieses notwendige Kriterium kann also grundsätzlich als erfüllt betrachtet werden. Es kann jedoch nicht gezeigt werden, dass die Unsicherheitsrückmeldung robust gegen die Bildung eines Cry Wolf Effektes ist. Die Ergebnisse beziehen sich jedoch in erster Linie auf den Vergleich zu einer Bedingung „keine Rückmeldung“. Es können aus den Ergebnissen von Studie eins noch keine Rückschlüsse gezogen werden, ob diese Effekte alleine der Charakteristik einer Unsicherheitsrückmeldung zuzuordnen sind, oder ob sie insbesondere für konventionell gestaltete Rückmeldungen auch gelten. Ebenso lässt sich aus den Ergebnissen nicht schließen, dass eine Unsicherheitsrückmeldung anders verarbeitet wird als ein normaler Alarm.

Als eine Konsequenz muss die Unsicherheitsrückmeldung zu anderen Alarmen verglichen werden. So könnte die Anfälligkeit für einen Cry Wolf Effekt bei einer Unsicherheitsrückmeldung zwar vorhanden, jedoch deutlich geringer sein als bei konventionellen Alarmen. In den Fokus der folgenden Versuche wird daher die Anfälligkeit gegenüber dem Cry Wolf Effekt und die Einflussfaktoren auf die Bildung eines solchen im Vergleich zwischen Unsicherheitsrückmeldung und anderen Rückmeldungsformen gelegt.

In den nachfolgenden Studien wird ein insgesamt realistischeres Verhältnis aus Situationen ohne Rückmeldung und Situationen mit Rückmeldung umgesetzt.

## 8 Studie 2

### 8.1 *Ziele Studie 2*

In Studie 1 konnten einige positive Effekte der Unsicherheitsrückmeldung beobachtet werden. So führt sie, verglichen mit keiner Rückmeldung, zu einer besseren Kontrollierbarkeit eines Automationsfehlers, wirkt sich puffernd auf die Entstehung von Miss-trust aus und führt insgesamt zu einer besseren Akzeptanz als bei einer Automation ohne Rückmeldesystem. Jedoch kann aus den Ergebnissen nicht geschlossen werden, ob diese Effekte auf die Charakteristik der Unsicherheitsrückmeldung zurückzuführen sind, oder ob diese nicht auch bei konventionellen Alarmen beobachtbar wären. Ebenso konnte eine Tendenz beobachtet werden, dass die Unsicherheitsrückmeldung unter wiederholter Darbietung von Rückmeldung ohne folgenden Automationsfehler einen Cry Wolf Effekt zeigt. Allerdings konnte die Größe des Cry Wolf Effektes nicht eindeutig quantifiziert werden, da keine Vergleichsmöglichkeit zu anderen Alarmstrategien existierte. Die Überprüfung der im Abschnitt Leithypothesen geäußerten Hypothese 3 „Unsicherheitsrückmeldung ist weniger anfällig für einen Cry Wolf Effekt, da keine feste Zusammenhangshypothese zwischen Rückmeldung und einer Konsequenz im Sinne einer sicheren Erwartung vorliegt“, konnte durch die in Studie 1 beobachteten Ergebnisse nicht hinreichend gestützt werden. Studie 2 soll hier weitere Einblicke ermöglichen. Zudem soll durch deutlich weniger Automationsfehler innerhalb einer Fahrt ein etwas realistischeres Szenario umgesetzt werden, als es in Studie 1 vorlag. Ebenso soll in der Gestaltung von Studie 2 der Kritikpunkt aus Studie 1 bezüglich des Verhältnisses aus Rückmeldungs- zu Nicht-Rückmeldungs- Situationen von fast 1:1 berücksichtigt werden. Ferner soll der Einfluss des Verhältnisses aus Nicht-Rückmeldungs- und Rückmeldungssituationen untersucht werden.

Im Folgenden werden die Teilziele aus Studie 2 und die damit verbundenen Forschungsfragen und Hypothesen erläutert.

**Ziel 1:** Untersuchung moderierender Einflüsse der Rückmeldungsart und der Rückmeldungshäufigkeit auf die Entstehung eines Cry Wolf Effektes anhand der Betrachtung der Kontrollierbarkeit.

**Forschungsfrage:** Wirkt sich die Art der Gestaltung der Rückmeldung unterschiedlich auf die Bildung einer Zusammenhangshypothese zwischen Rückmeldung und Automationsverhalten aus und beeinflusst die Art der Zusammenhangshypothese die Anfälligkeit für einen Cry Wolf Effekt?

**Hypothese:** Wenn die Unsicherheitsrückmeldung anders interpretiert wird als ein normaler Alarm ist die Neigung zu einem Cry Wolf Effekt bei gleich vielen erlebten Rückmeldungen ohne Automationsfehler geringer als bei einem normalen Alarm. Das heißt, wenn über die Fahrt nur Rückmeldungen ohne nachfolgenden Automationsfehler erlebt werden und in der letzten Fahrsituation tatsächlich ein Automationsfehler nach einer Rückmeldung folgt, dann ist bei Unsicherheitsrückmeldung die Erwartung eines Automationsfehlers höher und die Kontrollierbarkeit besser als in einer Bedingung mit normalem Alarm. Das heißt, es wird häufiger eine Kontrollreaktion gezeigt als in einer Bedingung normaler Alarm und die Reaktionszeit ist niedriger.

**Forschungsfrage:** Wird die Größe eines Cry Wolf Effektes von der Häufigkeit erlebter „falscher Alarme“ beeinflusst und wirkt sich die erlebte Häufigkeit von Rückmeldungen ohne Automationsfehler unterschiedlich auf die Rückmeldungsbedingungen Unsicherheit und normaler Alarm aus?

**Hypothese:** Wird das Ausbleiben eines Automationsfehlers nach einer Unsicherheitsrückmeldung tatsächlich nicht als irgendeine Art von Hypothesenevidenz angesehen, so hat die Häufigkeit der erlebten Kopplungen von Rückmeldung und keinem Automationsfehler keine Auswirkungen auf die Effektivität der Unsicherheitsrückmeldung in Kontrollierbarkeitssituationen. Bei einem normalen Alarm hingegen zeigt sich eine deutliche Verschlechterung der Kontrollierbarkeit eines Automationsfehlers von wenigen hin zu vielen vorher erlebten falschen Alarmen, da die Häufigkeit erlebter negativer Evidenz die ursprüngliche Erwartungshypothese beeinflusst.

**Ziel 2:** Untersuchung moderierender Einflüsse der Rückmeldungsart auf die Entstehung eines Cry Wolf Effektes anhand der Entwicklung der Nebenaufgabenleistung:

Das Ausmaß eines Cry Wolf Effektes sollte nicht nur über die Kontrollierbarkeit erfasst werden, sondern auch über die mögliche Veränderung der Aufmerksamkeitsverteilung zwischen einer Nebenaufgabe und der Fahraufgabe über die Fahrt hinweg. Als Indikator für die Aufmerksamkeitsverteilung wurde die Anzahl bearbeiteter Nebenaufgaben in einer Rückmeldesituation herangezogen. Nimmt die Erwartung, dass nach einer Rückmeldung keine negative Konsequenz folgt zu, dann müsste die Anzahl der bearbeiteten Nebenaufgaben in der Rückmeldungsperiode ansteigen.

**Forschungsfrage Kontrolle:** Ist die Nebenaufgabenleistung ein geeignetes Maß, um Rückmeldungseffektivität zu erfassen?

**Kontrollhypothese:** Wird einem Rückmeldesystem vertraut, dann wird in einer Rückmeldesituation der Fahraufgabe mehr Aufmerksamkeit geschenkt d.h. weniger Nebenaufgaben bearbeitet als in einem vergleichbaren Zeitraum einer Situation ohne Rückmeldung. Es gibt einen Unterschied in der Nebenaufgabenleistung zwischen Situationen mit Rückmeldung und Situationen ohne Rückmeldung.

**Forschungsfrage:** Gibt es einen Anstieg der Nebenaufgabenleistung in Rückmeldesituationen über die Fahrt und wie unterscheidet sich dieser zwischen den Rückmeldungsarten?

**Hypothese:** In einer Bedingung Unsicherheitsrückmeldung steigt die Nebenaufgabenleistung in Rückmeldesituationen nicht über die Fahrt an bzw., um Trainingseffekte zu berücksichtigen, nicht stärker an als die Nebenaufgaben-entwicklung in Situationen ohne Rückmeldung. In einer Bedingung mit normalem Alarm steigt die Nebenaufgabenleistung in Rückmeldesituationen über die Fahrt stärker an als in Situationen ohne Rückmeldung und deutlich stärker als in einer Unsicherheitsbedingung.

**Forschungsfrage:** Wirkt sich die erlebte Rückmeldungshäufigkeit unterschiedlich auf die Nebenaufgabenleistung in den beiden Rückmeldungsarten aus?

**Hypothese:** Wenn die Unsicherheitsrückmeldung nicht anfällig gegenüber einem Cry Wolf Effekt ist, dann müsste die durchschnittliche Nebenaufgabenleistung in Rückmeldesituationen unabhängig von der Anzahl erlebter „falscher Alarme“ gleichbleiben. D.h. innerhalb der Unsicherheitsbedingung darf es keinen Unterschied durch unterschiedlich häufig erlebte Rückmeldungen geben. In einer Bedingung normaler Alarm hingegen muss es einen Haupteffekt der Rückmeldungshäufigkeit geben. Insgesamt gibt es eine Interaktion aus Rückmeldungshäufigkeit und Rückmeldungsart.

**Ziel 3:** Untersuchung moderierender Einflüsse der Rückmeldungsart und der Rückmeldungshäufigkeit auf die Entstehung eines Cry Wolf Effektes anhand der Betrachtung der Entwicklung des situativen Vertrauens in die Automation. Zusätzlich zur allgemeinen Wirkung auf das Gesamtvertrauen in die Automation, soll auch die

kurzfristige, situative Auswirkung einer Rückmeldung auf das Vertrauen betrachtet werden.

**Forschungsfrage:** Wird durch eine Rückmeldung situativ das Vertrauen in die Automation verändert?

**Hypothese:** Die Rückmeldung passt das Vertrauen in die Automation situativ an. Werden keine oder nur wenige falsche Alarme erlebt, ist das situative Vertrauen nach Rückmeldungssituationen niedriger als zu einem gleichen Zeitpunkt nach einer Situation ohne Rückmeldung.

**Forschungsfrage:** Nimmt die Wirkung einer Rückmeldung auf die Anpassung des situativen Vertrauens mit zunehmender Anzahl erlebter falscher Alarme ab und wie ist diese Beziehung durch die Rückmeldungsart beeinflusst?

**Hypothese:** Es gibt eine Interaktion aus Rückmeldungshäufigkeit und Rückmeldungsart auf die Anpassung des situativen Vertrauens. Je mehr falsche Alarme erlebt werden, desto eher unterscheidet sich das situative Vertrauen zwischen einer Unsicherheitsrückmeldung und einem normalen Alarm. Bei wenigen falschen Alarmen ist das situative Vertrauen bei normalen Alarmen gleich der Unsicherheitsrückmeldung. Bei vielen erlebten falschen Alarmen ist in der Bedingung normaler Alarm das situative Vertrauen deutlich besser, d.h. weniger angepasst.

**Ziel 4:** Untersuchung der Auswirkungen von Rückmeldungsart und Rückmeldungshäufigkeit auf das Gesamtvertrauen in die Automation.

In Studie 1 wurde die puffernde Wirkung der Unsicherheitsrückmeldung auf die Entstehung von Misstrust beobachtet, wenn Unsicherheit mit keiner Rückmeldung verglichen wurde. In Studie 2 soll weiter untersucht werden, ob es ebenso rückmeldungsartbedingte Unterschiede in der Wirkung auf das Gesamtvertrauen gibt, wenn eine Unsicherheitsrückmeldung mit einem normalen Alarm verglichen wird.

**Forschungsfrage:** Wirken sich die Rückmeldungsarten Unsicherheitsrückmeldung und normaler Alarm und ihre postulierten unterschiedlichen Wirkungen auf die Entstehung eines Cry Wolf Effektes, auch unterschiedlich auf das Gesamtvertrauen in die Automation aus?

**Hypothese:** Es wird vermutet, dass nach dem Erleben eines Automationsfehlers, der aufgrund eines Cry Wolf Effektes nicht erwartet wurde, das Gesamtvertrauen in die Automation niedrig ist. Dabei gilt, je größer

der Cry Wolf Effekt, desto schlechter das allgemeine Vertrauen. In einer Bedingung mit Unsicherheitsrückmeldung wird unabhängig von der Anzahl erlebter falscher Alarmer in jeder Rückmeldesituation ein Automationsfehler für möglich gehalten. Das bedeutet, in der Bedingung Unsicherheit bleibt das situative Vertrauen unabhängig von der Anzahl erlebter falscher Alarmer gleich, in der Bedingung normaler Alarm hingegen nimmt das Gesamtvertrauen in die Automation mit zunehmender Anzahl erlebter falscher Alarmer ab.

**Ziel 5:** Untersuchung der Auswirkungen von Rückmeldungsart und Rückmeldungs-häufigkeit auf die Akzeptanz der Automation.

**Forschungsfrage:** Wird die Akzeptanz in die Automation durch einen Cry-Wolf Effekt beeinflusst?

**Hypothese:** Wird, aufgrund eines Cry Wolf Effektes, ein Automationsfehler weniger gut kontrolliert, so wirkt sich dies auf die Akzeptanz der Automation insgesamt aus. Wenn eine Unsicherheitsrückmeldung weniger anfällig für einen Cry Wolf Effekt ist, dann ist die Akzeptanz in der Unsicherheitsbedingung nach einer Fahrt mit Automationsfehler höher als in der Bedingung normaler Alarm.

## *8.2 Methode Studie 2*

Grundsätzlich baut das in Studie 2 verwendete Untersuchungsszenario wieder auf dem in Abschnitt „Basis-Untersuchungsparadigma“ beschriebenen Szenario auf. Die Versuchspersonen erleben, wie auch schon in Studie 1, für die Automation klare als auch unklare Situationen. Die Fahrsituationen sind die gleichen die auch in Studie 1 verwendet wurden (Abbildung 10, Abbildung 11 & Abbildung 14). In Studie 2 soll jedoch insgesamt ein etwas realistischeres Verhältnis aus Rückmeldungssituationen und Situationen ohne Rückmeldung umgesetzt werden, was insgesamt eine deutliche Verlängerung der Fahrt und entsprechend deutlich mehr Fahrsituationen erfordert.

So werden die Fahrsituationen aus Abbildung 10, Abbildung 11 & Abbildung 14, je nach Versuchsbedingung, zu insgesamt 41 Fahrsituationen pro Fahrt kombiniert. Im Gegensatz zu Studie 1, wird in Studie 2 die Fahrt nicht auf zwei Fahrtabschnitte aufgeteilt. Die Versuchspersonen durchfahren alle 41 Situationen in einer ununterbrochenen Fahrt.

Ein weiterer wesentlicher Unterschied zu Studie 1 ist, dass Versuchspersonen in der Fahrt nur einen Automationsfehler in der letzten Fahrsituation erleben. Die vorherigen 40 Fahrsituationen werden genutzt, um je nach Bedingung,

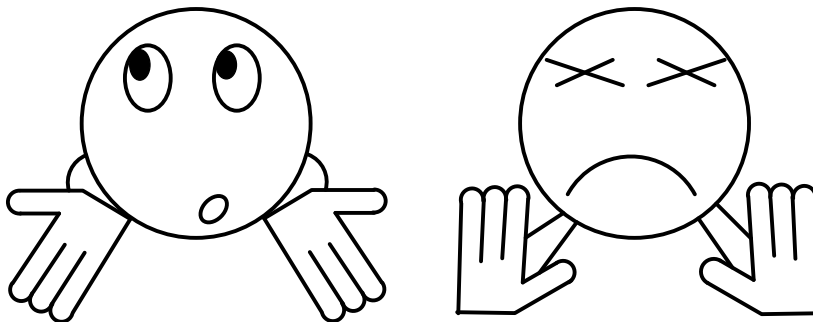
unterschiedlich häufige Erfahrungen von Rückmeldungen ohne Automationsfehler (falsche Alarmer) zu erzeugen. Auf Grundlage dieses Fahrscenarios werden, zur Umsetzung der einzelnen Studienziele, folgende Bedingungen bzw. Faktoren umgesetzt:

### 8.2.1 Versuchsdesign für Ziele 1-5:

Zur Umsetzung des Vergleichs der Unsicherheitsrückmeldung zu einem konventionell formulierten Alarm, wird ein Zwischengruppenfaktor Rückmeldungsart eingeführt. Dieser hat die Stufen:

1. Stufe: Unsicherheitsrückmeldung
2. Stufe: Normaler Alarm

Die unterschiedlichen Rückmeldungsarten Unsicherheitsrückmeldung und normaler Alarm, werden durch unterschiedliche Symbole und Instruktionen umgesetzt. Die Bedeutung der jeweiligen Rückmeldung wird vor Beginn der eigentlichen Untersuchung instruiert und die zugehörigen Symbole vorgeführt (siehe Abschnitt Umsetzung der Rückmeldungssituation).



**Abbildung 33: Symbol der Unsicherheitsrückmeldung (links) und des konventionellen Alarms (rechts).**

Zur Untersuchung der moderierenden Einflüsse der Häufigkeit von Rückmeldungen ohne Automationsfehler wird ein zweiter Zwischengruppenfaktor „Rückmeldungshäufigkeit“ eingeführt. Dieser hat die Stufen:

1. Stufe: niedrige Rückmeldungshäufigkeit (niedriges Cry Wolf Potenzial) - 4 Rückmeldungen ohne Automationsfehler in insgesamt 41 Fahrsituationen
2. Stufe: hohe Rückmeldungshäufigkeit (hohes Cry Wolf Potenzial) - 12 Rückmeldungen ohne Automationsfehler in insgesamt 41 Fahrsituationen

In der Bedingung niedrige Rückmeldungshäufigkeit (niedriges Cry Wolf Potenzial) wird nur in 4 von 41 Situationen, nämlich in der Situationsart „unklar“ (Fahrzeug halb

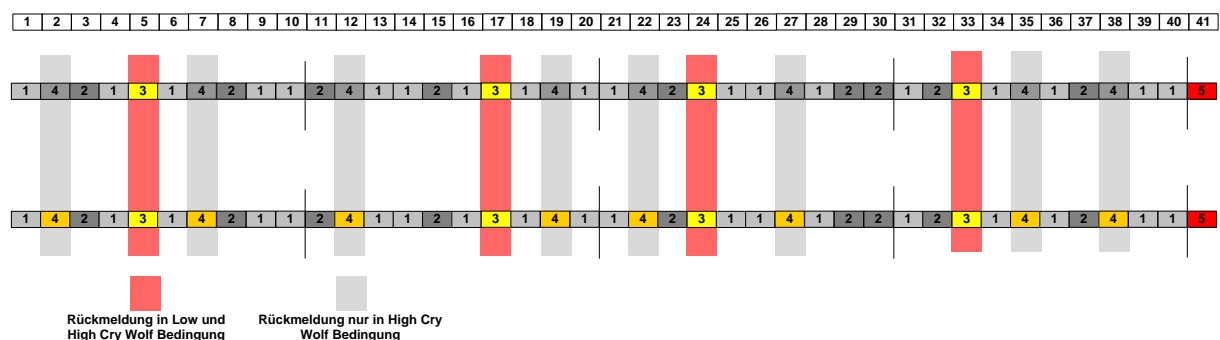


auf eigenem Fahrstreifen), eine Rückmeldung gegeben, auf die kein Automationsfehler folgt (Abbildung 34).

In den Fahrten der Bedingung hohe Rückmeldungshäufigkeit (hohes Cry Wolf Potenzial) wird zusätzlich in der Situationsart „unklar“ (Fahrzeug knapp auf Nachbarfahrstreifen) eine Rückmeldung gegeben, so dass hier in insgesamt 12 von 40 Situationen eine Rückmeldung erlebt wird auf die kein Automationsfehler folgt (Abbildung 34).

Zur Untersuchung der Entwicklung des situativen Vertrauens und der Nebenaufgabenleistung über die Fahrt wird zusätzlich ein Innergruppenfaktor „Wiederholung“ eingeführt. Durch Unterteilung der Fahrten in 4 gleich große Segmente – für die jeweils das durchschnittliche situative Vertrauen in den Rückmeldungssituationen ermittelt wird – hat der Wiederholungsfaktor quasi 4 Stufen, je Fahrtsegment eine Stufe.

Um einen direkten Vergleich mit der Bedingungen 4 und 12 Rückmeldungssituationen zu ermöglichen, werden die 12 Rückmeldungssituationen in der Bedingung hohe Rückmeldungshäufigkeit, durch Mittelung von je 3 Situationen, auf 4 Situationen, reduziert, (Abbildung 36).



**Abbildung 34: Abfolge der klaren und unklaren Situationen in den Fahrten der Bedingungen niedrige Rückmeldungshäufigkeit (obere Fahrt) und hohe Rückmeldungshäufigkeit (untere Fahrt). Der obere Zahlenstrahl stellt die fortlaufende Nummerierung der Situationen dar.**

## 8.2.2 Gesamtversuchsschema / Design

Aus der Faktorkombination der Zwischengruppenfaktoren Rückmeldungsart und Rückmeldungshäufigkeit ergibt sich zunächst ein 2 x 2 Zwischengruppen Design (Tabelle 15). Somit resultieren insgesamt 4 Versuchsbedingungen.

Für die Betrachtung der Entwicklung von Nebenaufgabenleistung und situativem Vertrauen über die Fahrt wird dieses Design um den Innergruppenfaktor „Wiederholung“ erweitert.

**Tabelle 15: Resultierender Versuchsplan aus der Kombination der Faktoren Rückmeldungsart und Rückmeldungshäufigkeit.**

		Cry Wolf Potenzial: (Between)	
		Niedrig	Hoch
Rückmeldungsart (Between)	Unsicherheit	y 1.1	y 1.2
	Normaler Alarm	y 2.1	y 2.2

### 8.2.3 Abhängige Variablen:

#### 8.2.3.1 Zu Ziel 1:

Untersuchung moderierender Einflüsse der Rückmeldungsart und der Rückmeldungshäufigkeit auf die Entstehung eines Cry Wolf Effektes anhand der Betrachtung der Kontrollierbarkeit. Dazu wurde die Kontrollierbarkeitsleistung in der Situation mit Automationsfehler (letzte Fahrsituation) betrachtet. Die Kontrollierbarkeit wurde operationalisiert durch:

- Reaktionshäufigkeit: Wie viele Versuchspersonen betätigten die Bremse
- Kollisionshäufigkeit – Wie viele Versuchspersonen kollidierten mit dem Vorderfahrzeug
- Reaktionszeit: Zeit zwischen dem Sichtbarwerden des Vorderfahrzeuges und der ersten Betätigung des Bremspedals

#### 8.2.3.2 Zu Ziel 2:

Untersuchung moderierender Einflüsse der Rückmeldungsart auf die Entstehung eines Cry Wolf Effektes anhand der Entwicklung der Nebenaufgabenleistung:

- Nebenaufgabenleistung in Form der Menge an bearbeiteten Nebenaufgaben in Rückmeldesituationen (gemessen vom Beginn der Rückmeldung bis zum Sichtbarwerden des Vorderfahrzeuges) – (siehe Abschnitt: „Messung des Cry Wolf Effektes“)
- Um nachvollziehen zu können, in wie weit ein Anstieg in der Nebenaufgabenleistung nicht auf einen Cry Wolf-, sondern auf einen Trainingseffekt in der Nebenaufgabenbearbeitung zurückzuführen ist, wurde als Referenzkriterium die Entwicklung der Nebenaufgabenleistung in Situationen ohne Rückmeldung verwendet und mit der Leistung in den Rückmeldungssituationen verglichen.

#### 8.2.3.3 Zu Ziel 3:

Untersuchung moderierender Einflüsse der Rückmeldungsart und der Rückmeldungshäufigkeit auf die Entstehung eines Cry Wolf Effektes anhand der Betrachtung der Entwicklung des situativen Vertrauens in die Automation.

- Situatives Vertrauen erfasst nach jeder Fahrsituation durch verbale Einschätzung auf einer 10stufigen Skala (siehe Erfassung des situativen Vertrauens).

#### **8.2.3.4 Zu Ziel 4:**

Untersuchung der Auswirkungen von Rückmeldungsart und Rückmeldungs-häufigkeit auf das Gesamtvertrauen in die Automation.

- Gesamtvertrauen erfasst nach der gesamten Fahrt, gemessen durch einen 5-dimensionalen Trust-Fragebogen (siehe Trustfragebögen).

#### **8.2.3.5 Zu Ziel 5:**

Untersuchung der Auswirkungen von Rückmeldungsart und Rückmeldungs-häufigkeit auf die Akzeptanz der Automation.

- Akzeptanz der Automation erfasst nach der gesamten Fahrt, gemessen anhand der Nutzungstendenz: (siehe Erfassung der Akzeptanz in die Automation)

### **8.2.4 Vorgehen**

Das Vorgehen war bis einschließlich des Trainings gleich dem Vorgehen in Studie 1 (siehe Methode Studie 1).

In den Versuchsbedingungen waren die Versuchspersonen wie in Studie 1 angehalten, während der Fahrt so viele Nebenaufgaben wie möglich zu bearbeiten, ohne dass das Fahren nach eigenem Ermessen unsicher wurde.

Um sowohl zur Nebenaufgabenbearbeitung als auch zum sicheren Fahren zu motivieren, wurde den Versuchspersonen in Aussicht gestellt, dass die Versuchsperson mit der besten Nebenaufgabenleistung einen Preis im Wert von etwa 10 Euro erhalten würde. Im Gegensatz dazu sollte eine sofortige Disqualifikation erfolgen, sobald eine kritische Fahrsituation herbeigeführt worden wäre (zur genaueren Erläuterung siehe Abschnitt Fahrerablenkung durch Nebenaufgaben).

Die Versuchspersonen wurden nach Situationen mit einem Fremdfahrzeug (unklar als auch klar) anhand einer 10 stufigen Skala nach ihrem situativen Vertrauen in die Automation befragt. In der 41. Situation der Fahrt trat ein Automationsfehler in der Form auf, dass die Automation hinter einem langsameren Fahrzeug nicht bremste, sondern die Versuchspersonen die Bremsung übernehmen mussten.

Nach dieser Kontrollierbarkeitssituation war die Fahrt vorbei. Die Versuchspersonen wurden gebeten Fragebögen zu bearbeiten, mit denen das Gesamtvertrauen in die Automation als auch die Akzeptanz erhoben wurden.

Danach erhielten die Versuchspersonen die Aufwandsentschädigung, wurden über den wahren Zweck der Untersuchung aufgeklärt und verabschiedet.

## 8.3 Ergebnisse Studie 2

### 8.3.1 Stichprobe

Die Versuchspersonen für Studie 2 wurden aus dem Probandenpool des DLR Instituts für Verkehrssystemtechnik akquiriert. Die 40 eingeladenen Versuchspersonen wurden nach dem in Abschnitt „Generelles Vorgehen: Stichproben“ beschriebenen Parallelisierungsschema auf die vier Versuchsgruppen aufgeteilt. Es wurden, nach den ebenfalls im Abschnitt „Generelles Vorgehen: Stichproben“ beschriebenen Kriterien, 7 Versuchspersonen von der Datenauswertung ausgeschlossen (Tabelle 16).

**Tabelle 16: Statistik der Substichproben aus Studie 2**

Gruppe	N	Männlich / weiblich	Alter Min / Max	Alter Mittelwert / Stdabw	Führerschein Min / Max	Führerschein Mittelwert / Stdabw
Unsicherheit Niedriges Cry Wolf Potenzial	9	5 / 4	20 / 53	32,4 / 11,9	2 / 32	12,6 / 11,0
Unsicherheit Hohes Cry Wolf Potenzial	9	6 / 3	18 / 65	37,0 / 17,9	1,5 / 47	19,0 / 17,6
Normaler Alarm Niedriges Cry Wolf Potenzial	7	4 / 3	21 / 61	36,0 / 26,0	4 / 43	18,0 / 16,2
Normaler Alarm Hohes Cry Wolf Potenzial	8	4 / 4	21 / 66	40,1 / 18,4	3 / 47	21,8 / 17,7

### 8.3.2 Ergebnisse der Kontrollierbarkeit von Automationsfehlern

Im Folgenden werden die Ergebnisse der Auswirkungen von Rückmeldungsart und Rückmeldungshäufigkeit auf die Reaktionshäufigkeit, die Reaktionszeit und die Kollisionshäufigkeit berichtet. Diese drei abhängigen Variablen dienen zur Erfassung der Qualität der Kontrollierbarkeit eines unangemessenen Automationsverhaltens.

#### 8.3.2.1 Reaktionshäufigkeit:

In einigen Fällen kam es vor, dass die Versuchspersonen in der Situation Rückmeldung und anschließender Automationsfehler, gar nicht auf das

Vorderfahrzeug reagierten und entsprechend nicht bremsen. Die Häufigkeiten sind in Tabelle 17 aufgeführt.

Da, bei einer ausbleibenden Reaktion, kein sinnvoller Parameter für eine Reaktionszeitanalyse bestimmt werden kann, werden die Versuchspersonen, die keine Reaktion zur Kollisionsvermeidung zeigten bei der weiteren Analyse der Reaktionszeit nicht berücksichtigt.

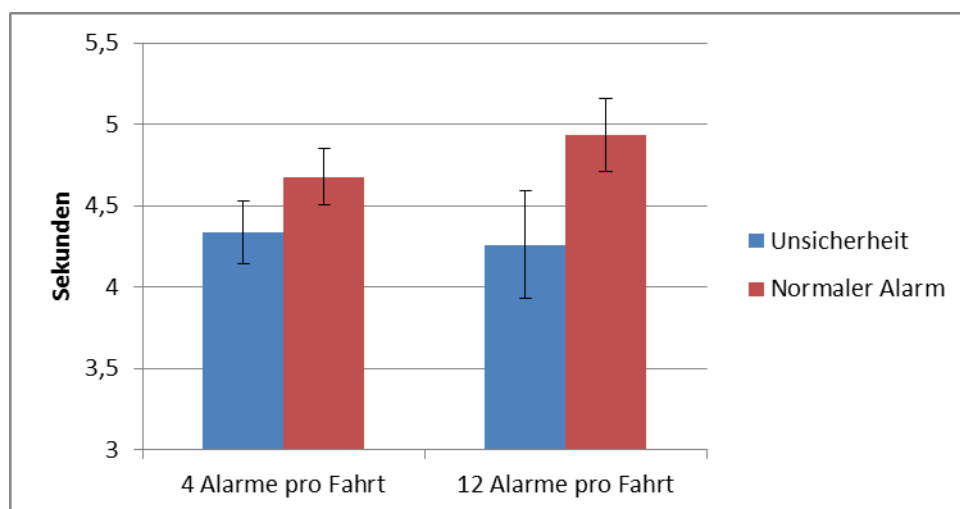
Bei der Analyse der Häufigkeiten derjenigen die keine Reaktion zeigten fällt auf, dass in der Unsicherheitsbedingung bei hohem Cry Wolf Potenzial, 2 Versuchspersonen keine Bremsreaktion beim Automationsfehler zeigten, in der Gruppe mit konventionellem Alarm nur jeweils eine Versuchsperson.

**Tabelle 17: Anzahl der Versuchspersonen die in der Kontrollierbarkeitssituation gar keine Reaktion zeigten – dahinter jeweils die Anzahl der Versuchspersonen in der Gruppe.**

	4 Rückmeldungen ohne Fehler / Fahrt	12 Rückmeldungen ohne Fehler / Fahrt
Unsicherheit	0 / 9	2 / 9
Normaler Alarm	1 / 7	1 / 8

### 8.3.2.2 Reaktionszeit

Alle Versuchspersonen, die eine Bremsreaktion zeigten, werden einer genaueren Analyse bezüglich der Qualität ihrer Reaktion unterzogen. Die Versuchspersonen, die nicht reagierten, werden von der Analyse ausgeschlossen, da für diese keine Reaktionszeit bestimmt werden kann.



**Abbildung 35: Bremsreaktionszeit in Kontrollierbarkeitssituation für beide Rückmeldungsgruppen und beide Rückmeldungshäufigkeiten.**

Auf deskriptiver Ebene werden zunächst die Unterschiede zwischen den Rückmeldungsarten betrachtet. Es zeigt sich augenscheinlich, dass die Unsicherheitsgruppe in beiden Bedingungen mit unterschiedlicher Rückmeldungshäufigkeit eine niedrigere Reaktionszeit aufweist als die Gruppe mit normalem Alarm (siehe Abbildung 35). Die zweifaktorielle Varianzanalyse zeigt, dass der Effekt des Faktors Rückmeldung auf die Reaktionszeit signifikant und von mittlerer Effektstärke ist. Aufgrund der geringen Stichprobengröße wird zur Absicherung ein Mann-Whitney U Test durchgeführt. Dieser kann das Ergebnis der ANOVA bestätigen (Tabelle 18).

**Tabelle 18: Ergebnisse der ANOVA & Mann-Whitney-U Test für die Faktoren Rückmeldung und Rückmeldungshäufigkeit**

Faktor	F	df	Sig.	Eta	Mann-Whitney-U-Test
Bremsreaktionszeit – Effekt der Rückmeldung	4,457	1/25	<b>.045</b>	,151	.029
Bremsreaktionszeit - Effekt der Rückmeldungshäufigkeit	,0139	1/25	.712	,005	--
Bremsreaktionszeit Interaktion Rückmeldung*Rückmeldungshäufigkeit	,475	1/25	.497	,019	--

Wird der Einfluss der erlebten Rückmeldungshäufigkeit betrachtet, so fällt bei Betrachtung der Abbildung 35 auf, dass es in der Bedingung Unsicherheitsrückmeldung kaum einen Unterschied in den Reaktionsparametern zwischen den beiden unterschiedlich hohen Rückmeldungshäufigkeiten gibt. In der Bedingung normaler Alarm gibt es hingegen augenscheinlich einen etwas deutlicher ausgeprägten Unterschied der Reaktionszeitmittelwerte. Die ANOVA kann aber keinen Interaktionseffekt aus Rückmeldungsart und Rückmeldungshäufigkeit zeigen (Tabelle 18).

### 8.3.2.3 Kollisionshäufigkeit

Tabelle 19 zeigt die Häufigkeit der Versuchspersonen, die trotz erfolgter Reaktion eine Kollision nicht vermeiden konnten. Die Versuchspersonen, die trotz Reaktion auch kollidierten, zeigen erwartungsgemäß auch innerhalb ihrer Gruppe die höchsten Reaktionszeiten. Allein aufgrund der Häufigkeiten, scheint es eher keinen Unterschied zwischen der Art der Rückmeldung zu geben, aber tendenziell einen Effekt der Rückmeldungshäufigkeit.

**Tabelle 19: Kollisionshäufigkeiten in der Kontrollierbarkeitssituation für die Versuchspersonen die eine Bremsreaktion zeigten. Angegeben ist die Anzahl der Kollisionen bezogen auf die Gesamtanzahl der Versuchspersonen mit Bremsreaktion in dieser Bedingung.**

	4 Rückmeldungen ohne Fehler / Fahrt	12 Rückmeldungen ohne Fehler / Fahrt
Unsicherheit	1 / 9	2 / 7
Normaler Alarm	0 / 6	2 / 7

Werden auch diejenigen Versuchspersonen mit in die Betrachtung der Kollisionshäufigkeit aufgenommen, welche gar keine Reaktion zeigten, so ist es der Effekt der Rückmeldungshäufigkeit auf die Kontrollierbarkeit eines Automationsfehlers deutlich augenscheinlicher, und zwar anscheinend unabhängig von der Art der Rückmeldung (Tabelle 20). Ein Ergebnis, das zunächst gegensätzlich zu den Ergebnissen der Reaktionszeitanalyse zu sein scheint.

**Tabelle 20: Anzahl der Versuchspersonen, die mit dem Vorderfahrzeug kollidierten, hinter dem Schrägstrich steht die Anzahl der Versuchspersonen dieser Gruppe (Anzahl nach Ausschluss s.o.)**

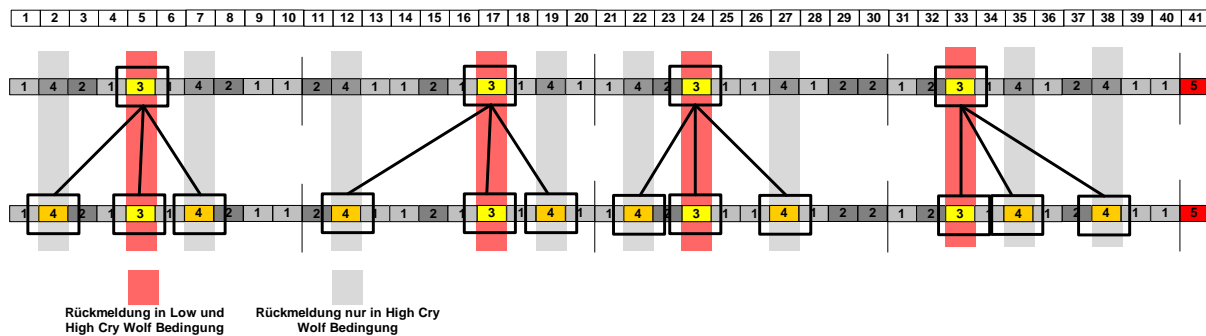
	4 Alarmer ohne Fehler pro Fahrt	12 Alarmer ohne Fehler pro Fahrt
Unsicherheit	1 / 9	4 / 9
Normaler Alarm	1 / 7	3 / 8

Aufgrund der ungleichmäßig besetzten Zellen ist es schwierig dieses Ergebnis anhand der absoluten Zahlen zu interpretieren. Auch die Interpretation des relativen Anteils der Kollisionen an der jeweiligen Zellgröße ist bei der geringen Fallzahl schwierig. Daher wird auch von der Durchführung eines Chi Quadrattests zu Häufigkeitsanalyse abgesehen.

### 8.3.3 Ergebnisse der Nebenaufgabenleistung

Im Folgenden werden die Ergebnisse der Auswirkungen von Rückmeldungsart und Rückmeldungshäufigkeit auf die Nebenaufgabenleistung berichtet.

Wegen der unterschiedlich häufigen Rückmeldungssituationen in den beiden Rückmeldungshäufigkeitsbedingungen, werden für die Bedingung mit 12 Rückmeldungssituationen jeweils drei aufeinanderfolgende Rückmeldungssituationen durch Bildung des Mittelwerts zu einem Messpunkt zusammengefasst. So ergeben sich in beiden Rückmeldungsbedingungen 4 Messpunkte die einen Vergleich mittels Messwiederholungs- ANOVA ermöglichen (Abbildung 36).



**Abbildung 36: Clusterung der Nebenaufgabenleistung von jeweils drei Rückmeldungssituationen in der Bedingung 12 Rückmeldungen pro Fahrt durch Bildung des Durchschnitts. Zweck war die Herstellung der Vergleichbarkeit der Nebenaufgabenleistung in der Bedingung 4 Alarmsituationen pro Fahrt mittels Messwiederholungs-Varianzanalyse.**

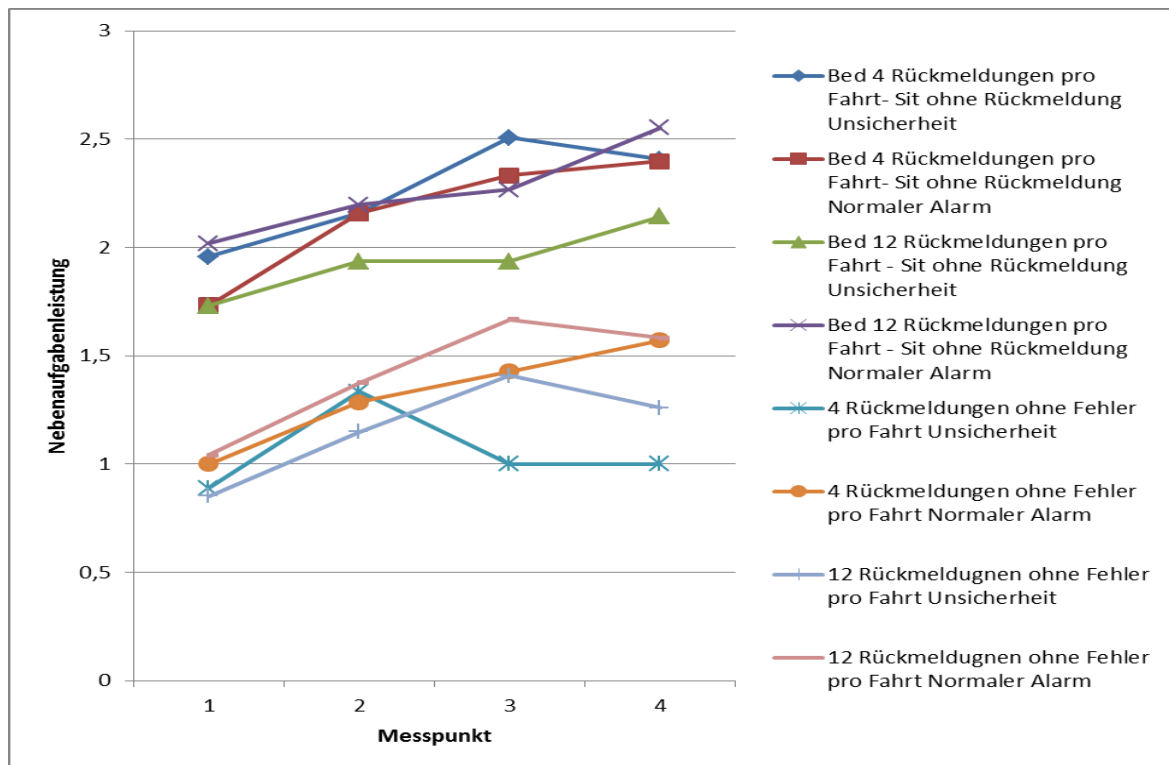
Zur Einschätzung des möglichen Einflusses eines Trainingseffektes, der durch das wiederholte Bearbeiten der Nebenaufgaben zustande kommen kann, wird zum Vergleich die Entwicklung der Nebenaufgabenleistung in den Fahrsituationen betrachtet, in denen keine Rückmeldung stattfand.

Die Unterscheidung Rückmeldungssituation und Nicht-Rückmeldungssituation bildet den Kontrollfaktor „Situationsklasse“.

Um, wie auch bei dem Vergleich der Rückmeldesituationen, auf 4 Vergleichspunkte zu kommen, wird ebenfalls die durchschnittliche Nebenaufgabenleistung über mehrere Situationen ohne Rückmeldung berechnet. Konkret bedeutet das, dass in der Bedingung mit 4 Rückmeldungssituationen je Fahrt die durchschnittliche Nebenaufgabenleistung von jeweils 9 Situationen ohne Rückmeldung ermittelt wird, um auf insgesamt 4 gemittelte Messpunkte zu kommen. Entsprechend wird in der Bedingung 12 Rückmeldungssituationen pro Fahrt der Durchschnitt der Nebenaufgabenleistung von 7 Situationen ohne Rückmeldung gebildet, um auf die 4 Messpunkte zu kommen (Abbildung 37).

Die ermittelten 4 Messpunkte fließen in eine Mixed Models Varianzanalyse ein, mit den 4 Messpunkten als Messwiederholungsfaktor, sowie der Rückmeldungsart, der Rückmeldungshäufigkeit und der Situationsklasse als Zwischensubjektfaktoren.





**Abbildung 37: Entwicklung der Nebenaufgabenleistung über die Fahrt in Situationen mit und ohne Rückmeldung. Zur besseren Vergleichbarkeit wurde die Nebenaufgabenleistung durch Durchschnittsbildung zu jeweils vier Vergleichspunkten geclustert.**

Deskriptiv ergibt sich ein deutlicher Unterschied zwischen den Situationen mit und ohne Rückmeldung. In Situationen mit Rückmeldung werden durchschnittlich deutlich weniger Nebenaufgaben bearbeitet als in den Situationen ohne Rückmeldung (Abbildung 37). Dies gilt für beide Rückmeldungsarten.

Innerhalb der Situationsklasse Rückmeldung ohne Automationsfehler scheint es einen geringen systematischen Unterschied zwischen den Rückmeldungsarten zu geben. Die durchschnittliche Nebenaufgabenleistung ist in den Rückmeldungssituationen mit normalem Alarm etwas höher als in der Bedingung Unsicherheitsrückmeldung.

**Tabelle 21: Ergebnisse der Messwiederholungs-ANOVA für den Innersubjektfaktor Wiederholung (Entwicklung über die 4 Messpunkte) und die Zwischensubjektfaktoren Situationsklasse (Rückmeldung vs. Keine Rückmeldung), Rückmeldungsart (Unsicherheit vs. Normaler Alarm) & Rückmeldungshäufigkeit (4 Rückmeldungen vs. 12 Rückmeldungen)**

<b>Faktor</b>	<b>F</b>	<b>df</b>	<b>Sig.</b>	<b>Eta</b>
<b>Wiederholung (Innersubjektfaktor)</b>	6,875	3 / 171	<b>,000</b>	,108
<b>Situationsklasse (Zwischensubjektfaktor)</b>	27,512	1 / 57	<b>,000</b>	,326
<b>Rückmeldungsart (Zwischensubjektfaktor)</b>	,932	1 / 57	,338	,016
<b>Rückmeldungshäufigkeit (Zwischensubjektfaktor)</b>	,093	2 / 57	,911	,003
<b>Wiederholung * Rückmeldungsart</b>	,801	3 / 171	,495	,014
<b>Wiederholung * Rückmeldungshäufigkeit</b>	,842	3 / 171	,539	,029
<b>Wiederholung * Situationsklasse</b>	,880	3 / 171	,453	,015
<b>Wiederholung * Rückmeldungsart* Rückmeldungshäufigkeit</b>	,211	3 / 171	,888	,004

Wird die Entwicklung der Nebenaufgabenleistung über die vier Messpunkte betrachtet, so steigt diese in fast allen Bedingungen über die vier Situationen hinweg an (Ausnahme Bedingung Rückmeldungssituation / Unsicherheit / 4 Rückmeldungen pro Fahrt). Es gibt einen deutlichen Unterschied in der Nebenaufgabenleistung zwischen Situationen mit und ohne Rückmeldung.

Die Varianzanalyse bestätigt den offensichtlichen Effekt der Situationsklasse (Tabelle 21). Für die Faktoren Rückmeldungsart und Rückmeldungshäufigkeit ergibt sich jedoch kein Haupteffekt (Tabelle 21).

Da in den Situationen mit Rückmeldung augenscheinlich ein gewisser Unterschied zwischen den Rückmeldungsarten vorliegt, werden mögliche Effekte der Rückmeldungsart mit einem Mann-Whitney U Test überprüft. Dafür wird zunächst je Versuchsperson der Durchschnitt der bearbeiteten Nebenaufgaben in den Alarmsituationen berechnet. Auch der U-Test kann keinen signifikanten Unterschied zwischen den Rückmeldungsarten zeigen (Tabelle 22).

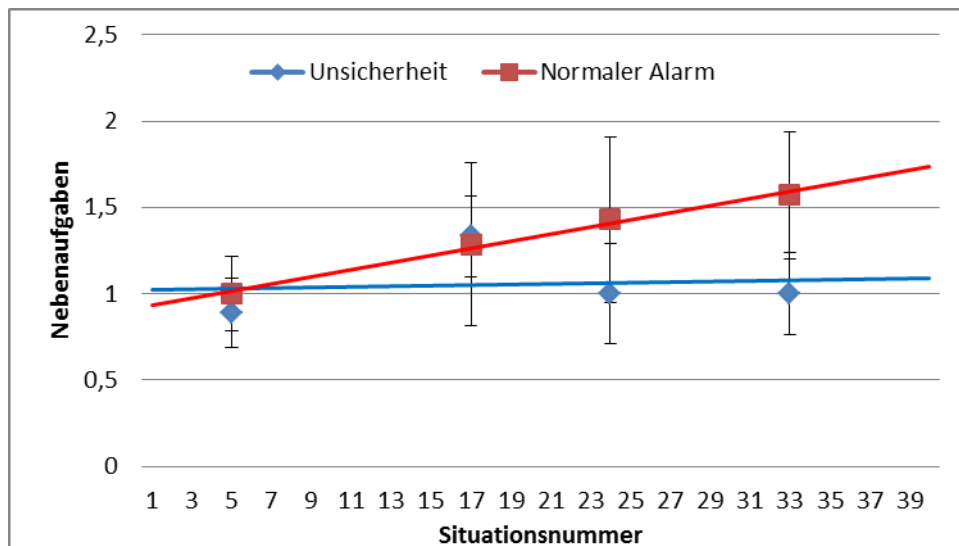
**Tabelle 22: Ergebnisse Mann-Whitney U-Test für den Unterschied in der Nebenaufgabenleistung in Situationen mit Rückmeldung zwischen Unsicherheit und normalem Alarm sowohl in der Bedingung 4 Rückmeldungen / Fahrt und 12 Rückmeldungen / Fahrt.**

	p.
Unterschied Unsicherheit – Normaler Alarm bei 4 Rückmeldungen / Fahrt	0.5229
Unterschied Unsicherheit – Normaler Alarm bei 12 Rückmeldungen / Fahrt	0.4694

Die Varianzanalyse zeigt einen Haupteffekt des Faktors Wiederholung (Tabelle 21). Es ergeben sich jedoch keine Interaktionseffekte aus dem wiederholten Erleben von Rückmeldungssituationen und der Art der Rückmeldung (normaler Alarm & Unsicherheitsrückmeldung). Ebenso kann kein Interaktionseffekt aus der Wiederholung und der Situationsklasse oder der Rückmeldungshäufigkeit festgestellt werden (Tabelle 21).

Die Nebenaufgabenleistung steigt, anscheinend unabhängig von der Art der Rückmeldung und der erlebten Rückmeldungshäufigkeit, über die in der Fahrt erlebten Situationen hinweg an (Abbildung 37).

Zwecks besserer Interpretierbarkeit der Nebenaufgabenentwicklung werden zwei weitere Analysen der Nebenaufgabenentwicklung über die Rückmeldungssituationen, jeweils getrennt für die beiden Rückmeldungshäufigkeiten, durchgeführt.



**Abbildung 38: Entwicklung der Nebenaufgabenleistung in der Bedingung niedrige Rückmeldungshäufigkeit. Die Linien stellen lineare Regressionsgraden dar.**

Wird die Bedingung mit niedriger Rückmeldungshäufigkeit alleine betrachtet (Abbildung 38), zeigt sich, bei den Situationen mit Rückmeldung für die Bedingung

normaler Alarm ein ansteigender Trend über die 4 Situationen, jedoch nicht für die Bedingung Unsicherheit.

Die Messwiederholungs-ANOVA kann jedoch keinen Interaktionseffekt aus der wiederholt erlebten Darbietung einer Rückmeldung ohne Automationsfehler und der Rückmeldungsart zeigen (Tabelle 23).

Auch ein Friedmann Test zeigt in der Bedingung Unsicherheit keinen signifikanten Effekt des Faktors Wiederholung auf die Nebenaufgabenentwicklung, aber auch in der Bedingung normaler Alarm kann kein signifikanter Anstieg der Nebenaufgabenleistung gezeigt werden (Tabelle 24).

**Tabelle 23: Ergebnisse der Mixed models ANOVA für die Bedingung niedrige Rückmeldungshäufigkeit (4 Rückmeldungssituationen pro Fahrt)**

Faktor	F	df	Sig.	Eta
<b>Wiederholung (Innersubjektfaktor)</b>	1,040	3 / 27	,391	,104
<b>Rückmeldungsart (Zwischensubjektfaktor)</b>	,190	3 / 27	,673	,021
<b>Wiederholung * Rückmeldungsart</b>	,485	3 / 27	,695	,051

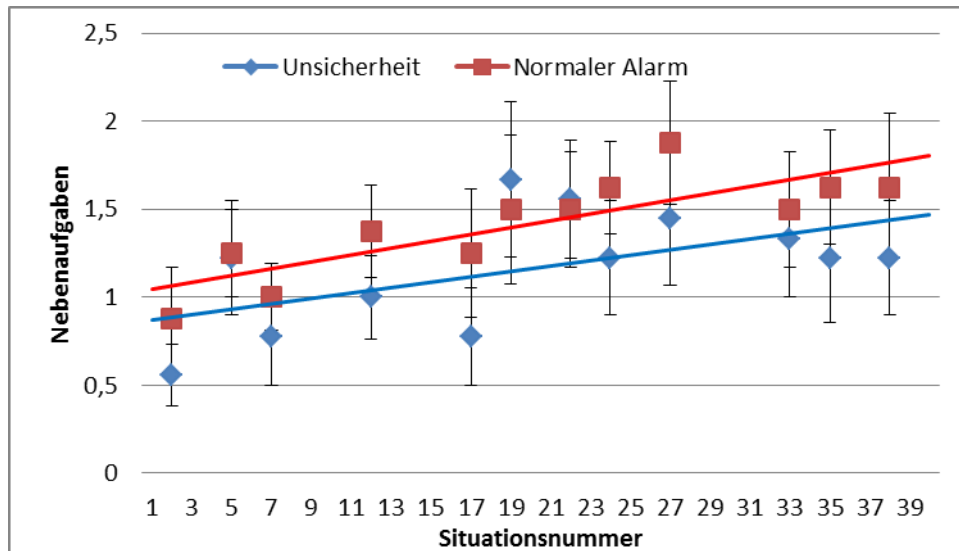
**Tabelle 24: Ergebnisse Friedmann-Test der Nebenaufgabenentwicklung über 4 Rückmeldungssituationen pro Fahrt**

	X2	df	P
<b>Unsicherheit</b>	1,9412	3	.5847
<b>Normaler Alarm</b>	2.1818	3	.5355

In der Bedingung mit hoher Rückmeldungshäufigkeit steigt die Nebenaufgabenleistung sowohl in der Bedingung normaler Alarm als auch in der Bedingung Unsicherheit über die Fahrt hinweg gleich an, es ist kein augenscheinlicher Unterschied zwischen den Rückmeldungsarten zu erkennen. Die Messwiederholungs-ANOVA kann insgesamt den Anstieg der Nebenaufgabenleistung als Haupteffekt bestätigen. Es ergeben sich aber weder ein Haupteffekt für die Rückmeldungsart noch ein Interaktionseffekt aus Rückmeldungsart und Wiederholung. Ein Friedmantest kann allerdings den in der ANOVA gefundenen Haupteffekt des Wiederholungsfaktors auf die Nebenaufgabenentwicklung nicht bestätigen (Tabelle 25).

**Tabelle 25: Ergebnisse Friedmann-Test der Nebenaufgabenentwicklung über 12 Rückmeldungssituationen pro Fahrt**

	X2	df	P
<b>Unsicherheit</b>	17,0359	11	.1068
<b>Normaler Alarm</b>	15,805	11	.1485



**Abbildung 39. Entwicklung der Nebenaufgabenleistung in der Bedingung hohe Rückmeldungshäufigkeit. Die Linien stellen lineare Regressionsgraden dar.**

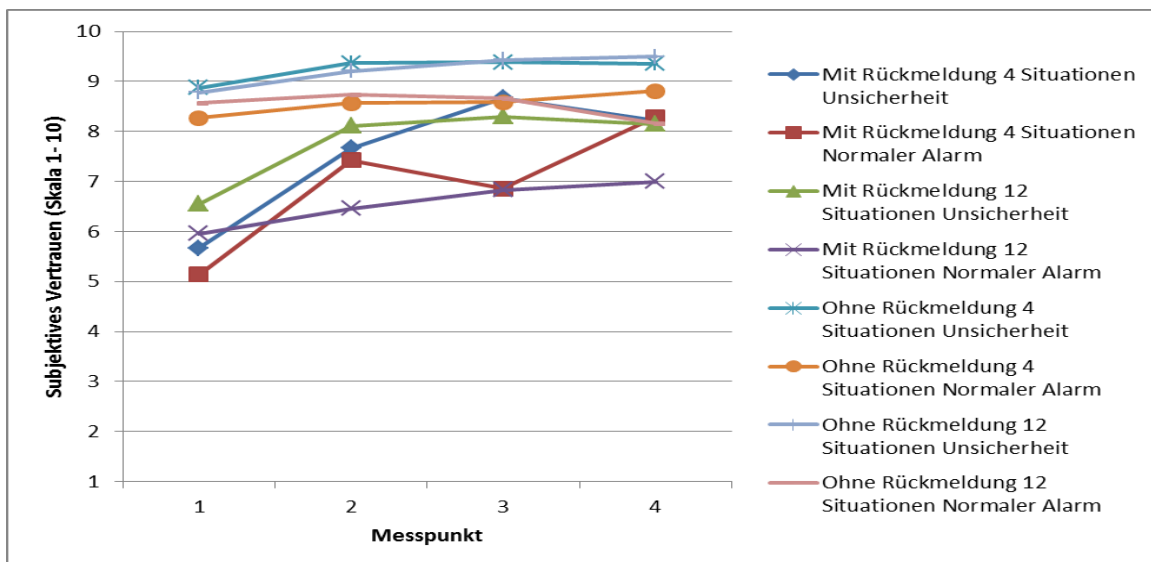
**Tabelle 26: Ergebnisse der Mixed models ANOVA für die Bedingung hohe Rückmeldungshäufigkeit (12 Rückmeldungen pro Fahrt)**

Faktor	F	df	Sig.	Eta
<b>Wiederholung (Innersubjektfaktor)</b>	5,642	3 / 45	<b>,002</b>	,273
<b>Rückmeldungsart (Zwischensubjektfaktor)</b>	,598	1 / 15	,451	,038
<b>Wiederholung * Rückmeldungsart</b>	,070	3 / 45	,976	,005

### 8.3.4 Ergebnisse situatives Vertrauen

Im Folgenden werden die Ergebnisse der Auswirkungen von Rückmeldungsart und Rückmeldungshäufigkeit auf das situative Vertrauen in die Automation berichtet. Um alle Faktoren in einer Mixed-models Varianzanalyse gemeinsam betrachten zu können, müssen zunächst, wie bereits in der Auswertung der Nebenaufgabenentwicklung, die situativen Vertrauensmessungen der Bedingung mit 12 Rückmeldungen pro Fahrt auf vier Messpunkte aggregiert werden. Auch die Messungen des situativen Vertrauens in der Situationsklasse Nicht-Rückmeldungssituationen müssen auf 4 Messpunkte zusammengefasst werden.

Deskriptiv (Abbildung 40) ist das situative Vertrauen in den Situationen ohne Rückmeldung durchgehend auf hohem Niveau und scheint sich über die Fahrt nicht deutlich zu erhöhen. Das situative Vertrauen in den Situationen ohne Rückmeldung ist in der Unsicherheitsbedingung insgesamt etwas höher als in der Bedingung normaler Alarm. Dies scheint unabhängig von der erlebten Rückmeldungshäufigkeit zu sein. In Rückmeldungssituationen ist das situative Vertrauen insgesamt niedriger als in Situationen ohne Rückmeldung, jedoch ist eine deutliche Entwicklung des situativen Vertrauens über die Fahrt beobachtbar. Zu Beginn der Fahrt ist das situative Vertrauen nach Rückmeldungssituationen noch relativ niedrig, im Vergleich zu Situationen ohne Rückmeldung. Es steigt aber über die erlebten Alarmsituationen ohne Automationsfehler hinweg tendenziell an. Dieser Anstieg ist dabei insgesamt in der Bedingung Unsicherheitsrückmeldung ausgeprägter als in der Bedingung normaler Alarm.



**Abbildung 40: Entwicklung des situativen Vertrauens über die Fahrt bei hoher und niedriger Rückmeldungshäufigkeit in Situationen mit und ohne Rückmeldung jeweils in den Bedingungen Unsicherheitsrückmeldung und Normaler Alarm.**

Eine Mixed Models Varianzanalyse mit der wiederholten Messung über die Fahrt als Innersubjektfaktor und Rückmeldungsart, Rückmeldungshäufigkeit und Situationsklasse als Zwischensubjektfaktoren (Tabelle 27), bestätigt einen Haupteffekt des Zwischensubjektfaktors Situationsklasse. In, bzw. kurz nach Rückmeldungssituationen ist das situative Vertrauen in die Automation insgesamt niedriger als nach Situationen ohne Rückmeldung.

Die Rückmeldungsart hat ebenfalls einen signifikanten Haupteffekt, das situative Vertrauen ist in der Unsicherheitsbedingung insgesamt höher als in der Bedingung normaler Alarm.

Die Rückmeldungshäufigkeit scheint keine Auswirkungen zu haben, hier kann, wie auf deskriptiver Ebene bereits vermutet, insgesamt kein Haupteffekt gezeigt werden.

Zu dem beobachteten Anstieg des situativen Vertrauens über die 4 Vergleichspunkte, zeigt die Varianzanalyse einen signifikanten Haupteffekt des Wiederholungsfaktors. Ebenso zeigt sie einen Interaktionseffekt aus der Wiederholung und der Situationsklasse. Der Anstieg ist dabei in den Alarmsituationen deutlicher als in der Bedingung Nicht-Alarmsituationen. Dabei liegt das situative Vertrauen in den Alarmsituationen zunächst deutlich unter dem Niveau der Nicht-Alarmsituationen. Es schließt dann aber über die Fahrt fast auf das Niveau der Nicht-Alarmsituationen auf. Zunächst vertrauen Versuchspersonen der Automation nach Alarmsituationen nicht, im Laufe der Fahrt wird der Automation in Alarmsituationen aber zunehmend vertraut.

Die Rückmeldungshäufigkeit hat keinen Haupteffekt, jedoch ist ein tendenzieller Interaktionseffekt der Faktoren Rückmeldungshäufigkeit und Wiederholung erkennbar.

**Tabelle 27: Ergebnisse der Mixed Models Varianzanalyse des situativen Vertrauens für die wiederholte Messung über die Fahrt als Innersubjektfaktor und Rückmeldungsart, Rückmeldungshäufigkeit und Situationsklasse als Zwischensubjektfaktoren.**

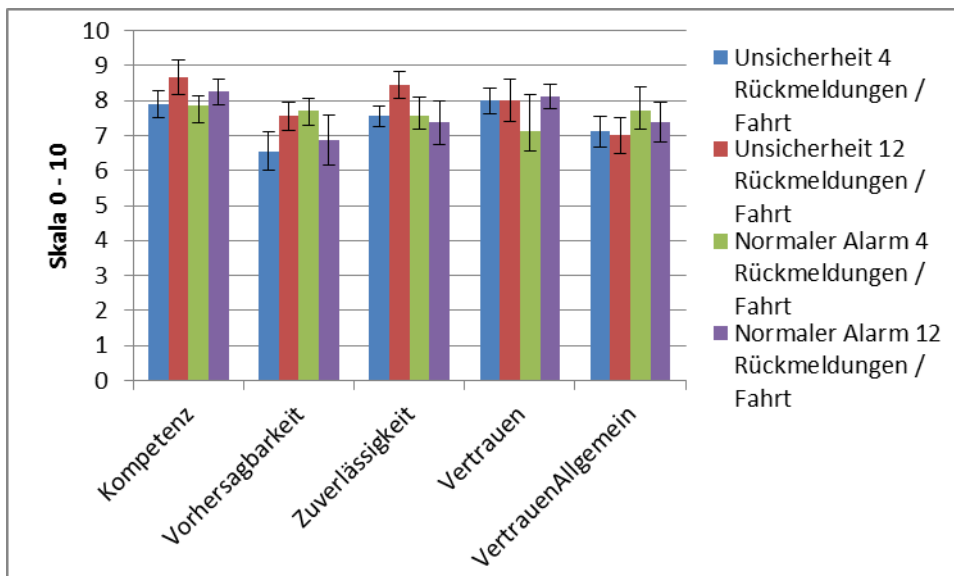
Faktor	F	df	Sig.	Eta
<b>Wiederholung (Innersubjektfaktor)</b>	19,33	1,69 / 174	<b>.000</b>	,250
<b>Situationsklasse (Zwischensubjektfaktor)</b>	21,43	1 / 58	<b>.000</b>	,270
<b>Rückmeldungsart (Zwischensubjektfaktor)</b>	4,90	1 / 58	<b>.031</b>	,078
<b>Rückmeldungshäufigkeit (Zwischensubjektfaktor)</b>	,018	1 / 58	.895	,000
<b>Wiederholung * Rückmeldungsart</b>	1,41	1,69 / 174	.249	,024
<b>Wiederholung * Rückmeldungshäufigkeit</b>	2,39	1,69 / 174	.105	,040
<b>Wiederholung * Situationsklasse</b>	9,21	1,69 / 174	<b>.000</b>	,137
<b>Wiederholung * Rückmeldungsart* Rückmeldungshäufigkeit</b>	1,29	1,69 / 174	.275	,022

### 8.3.5 Ergebnisse Gesamtvertrauen

Im Folgenden werden die Ergebnisse der Auswirkungen von Rückmeldungsart und Rückmeldungshäufigkeit auf das Gesamtvertrauen in die Automation berichtet.

Bei Betrachtung der Abbildung 41 fällt zunächst auf, dass sich das Vertrauen in die Automation in allen Trust-Dimensionen, unabhängig von der Rückmeldungsart und der Rückmeldungshäufigkeit, in einem recht hohen Bereich bewegt. Die Versuchspersonen vertrauen der Automation also, trotz der gemachten Erfahrungen mit Rückmeldungen ohne Automationsfehler und dem erlebten Automationsfehler, recht ausgeprägt. Es ergeben sich keine offensichtlichen Unterschiede zwischen den Rückmeldungsarten und auch nicht zwischen den Rückmeldungshäufigkeiten. Die multivariate Varianzanalyse kann dieses Bild stützen.

Es zeigt sich weder ein Haupteffekt der Rückmeldung ( $F_{,784} df(6/24) p = .591$  partielles  $\eta^2 = ,164$ ) noch ein Haupteffekt der Rückmeldungshäufigkeit ( $F_{,878} df(6/24) p = .525$  partielles  $\eta^2 = ,180$ ). Ebenso kann auf den einzelnen Trust-Dimensionen kein Effekt der Rückmeldungsart, kein Effekt der Rückmeldungshäufigkeit und auch kein Interaktionseffekt beobachtet werden (siehe Tabelle 28).



**Abbildung 41: Mittelwerte und Standardfehler des Gesamtvertrauens in die Automation nach Ende der Fahrten. Dargestellt sind die Werte für die beiden Rückmeldungsarten bei unterschiedlicher Fehlerhäufigkeit jeweils für die 5 Dimensionen des Trust-Fragebogens.**

Weder die hier betrachteten Rückmeldungsarten noch die erlebte Rückmeldungshäufigkeit, scheinen eine besondere Auswirkung auf das Gesamtvertrauen in die Automation zu haben.

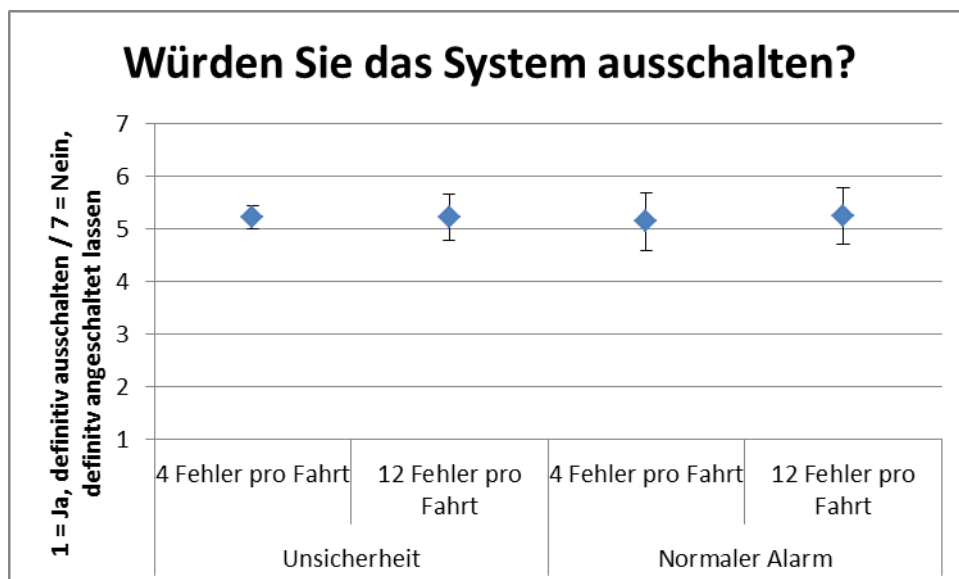


**Tabelle 28: Ergebnisse der MANOVA für die einzelnen Trust-Dimensionen für die Faktoren Rückmeldung und Rückmeldungshäufigkeit.**

	Trustdimension	df	F	Sig.	Partielles Eta
Rückmeldung	Kompetenz	1, 29	.303	.586	,010
	Vorhersagbarkeit	1, 29	.197	.661	,007
	Zuverlässigkeit	1, 29	1,313	.261	,043
	Vertrauen	1, 29	.362	.552	,012
	Vertrauen allgemein	1, 29	.784	.383	,026
Rückmeldungs- häufigkeit	Kompetenz	1, 29	2,066	.161	,067
	Vorhersagbarkeit	1, 29	.022	.883	,001
	Zuverlässigkeit	1, 29	.567	.457	,019
	Vertrauen	1, 29	.651	.426	,022
	Vertrauen allgemein	1, 29	.166	.686	,006

### 8.3.6 Ergebnisse Akzeptanz der Automation

Nach der Durchführung der Fahrt wurden die Versuchspersonen zur Akzeptanz des Gesamtsystems befragt. Die Versuchspersonen wurden gefragt, ob sie die Automation aufgrund der gemachten Erfahrungen ausschalten würden. Es ergaben sich augenscheinlich keine Unterschiede zwischen den Rückmeldungsarten und auch nicht zwischen den unterschiedlichen Rückmeldungshäufigkeiten (Abbildung 42).



**Abbildung 42: Ergebnisse der Beantwortung der Frage, ob die Versuchspersonen die Automation nach der gemachten Erfahrung ausschalten würden. Zu sehen sind die Mittelwerte mit den Standardfehlern des Mittelwertes.**

## 8.4 Diskussion Studie 2

In Studie 2 sollte untersucht werden, ob eine Unsicherheitsrückmeldung tatsächlich anders erlebt und verhaltenswirksam wird als ein normaler Alarm. Dazu wurde eine Unsicherheitsrückmeldung im Vergleich zu einem konventionell gestalteten Alarm betrachtet. Zur Überprüfung der These, dass die Wahrscheinlichkeit der Entstehung eines Cry Wolf Effektes und die Schwere dessen, eine Funktion der Anzahl der in einer Sequenz erlebten negativen Evidenzen ist, wurde als zweiter Faktor die Rückmeldungshäufigkeit der erlebten Rückmeldungen ohne Automationsfehler eingeführt. Diese wurde in zwei Stufen, hoch und niedrig variiert. Als dritter Faktor wurde die Entwicklung über die Fahrsituationen betrachtet, dies aber immer nur in Abhängigkeit der beiden Zwischengruppenfaktoren (Interaktionen).

Die in Abschnitt Leithypothesen formulierten Haupthypothesen sollten in Studie 2 durch die Bildung mehrerer Teilziele näher operationalisiert werden. Der Diskussionsteil ist daher nach den einzelnen, im Methodenabschnitt zu Studie 2 formulierten, Zielen gegliedert.

Bevor die unterschiedlichen Ergebnisse diskutiert werden, muss zunächst erwähnt werden, dass nicht alle Versuchspersonen in die Analyse einfließen. Einige der Versuchspersonen reagierten in den Rückmeldesituationen ohne Automationsfehler mit einer Bremsung und bekamen dadurch nicht mit, dass gar kein Automationsfehler stattfand. Da sich der Faktor Rückmeldungshäufigkeit auf die Anzahl erlebter Rückmeldungen ohne Automationsfehler bezog, konnte für diese Versuchspersonen die eigentliche Bedingung nicht hergestellt werden. Die angestrebte Zahl von 10 Versuchspersonen pro Gruppe konnte so nicht erreicht werden und die Teilstichproben waren zum Teil nur mit relativ wenigen Versuchspersonen besetzt ( $N = 9$  bis  $N = 7$ ). Es wurden daher in einigen Fällen neben den parametrischen Testverfahren, non-parametrische Tests genutzt. Die Ergebnisse müssen daher unter diesen Einschränkungen betrachtet werden.

**Ziel 1:** Untersuchung moderierender Einflüsse der Rückmeldungsart und der Rückmeldungshäufigkeit auf die Entstehung eines Cry Wolf Effektes anhand der Betrachtung der Kontrollierbarkeit.

Bei der Auswertung der Reaktions- und Kollisionshäufigkeit in der Situation mit Automationsfehlern fiel auf, dass es zwischen Unsicherheitsrückmeldung und dem normalen Alarm nur marginale und statistisch nicht absicherbare Unterschiede gibt. Es reagieren in beiden Rückmeldungsarten ähnlich viele Versuchspersonen gar nicht. Wenn eine Reaktion stattfindet, kollidierten in beiden

Rückmeldungsbedingungen trotzdem ähnlich viele Versuchspersonen mit dem Vorderfahrzeug.

Es scheint, unabhängig von der Rückmeldungsart, einen augenscheinlichen Einfluss der erlebten Rückmeldungshäufigkeit zu geben. D.h. bei niedriger Rückmeldungshäufigkeit werden deutlich weniger Kollisionen als bei hoher Rückmeldungshäufigkeit beobachtet.

Anscheinend gibt es unabhängig von der Rückmeldungsart eine Tendenz zu einem Cry Wolf Effekt. Dies spricht zunächst gegen die Hypothese, dass eine Unsicherheitsrückmeldung anders interpretiert wird als ein normaler Alarm.

Bei der Reaktionszeitanalyse zeigen die Versuchspersonen in den Unsicherheitsbedingungen eine insgesamt bessere Reaktionszeit als in der Bedingung mit normalem Alarm. Dies äußert sich allerdings nur in einem eher kleinen Effekt. Auch gibt es innerhalb der Unsicherheitsbedingungen keinen Reaktionszeitunterschied zwischen niedriger und hoher Rückmeldungshäufigkeit. Bei einem normalen Alarm ist, deskriptiv betrachtet, die Reaktionszeit bei hoher Rückmeldungshäufigkeit höher als bei niedriger Rückmeldungshäufigkeit. Es kann jedoch kein signifikanter Interaktionseffekt aus Rückmeldungsart und Rückmeldungshäufigkeit festgestellt werden.

Die vorliegenden Daten der Kollisionshäufigkeiten und der Reaktionsqualität scheinen zunächst widersprüchlich zu sein. Die Ergebnisse der Kollisionshäufigkeiten weisen darauf hin, dass es keinen Unterschied zwischen den Rückmeldungsarten bezüglich der Anfälligkeit für einen Cry Wolf Effekt gibt. Die Ergebnisse der Reaktionsanalyse stützten jedoch eher die Hypothese, dass es einen systematischen Unterschied gibt. Obwohl die Reaktionszeit in der Unsicherheitsbedingung bei niedriger und hoher Rückmeldungshäufigkeit ähnlich ist, kollidieren deutlich mehr Versuchspersonen in der Bedingung mit hoher Rückmeldungshäufigkeit. Eine mögliche Erklärung ist die, dass die Unsicherheitsrückmeldung nicht bei allen Versuchspersonen ähnlich wirkt. Ein Hinweis darauf ist der größere Standardfehler in der Bedingung hoher Rückmeldungshäufigkeit. Bei einigen Versuchspersonen wirkt die Unsicherheitsrückmeldung anscheinend deutlich effizienter als bei Anderen. Ein möglicher Grund für diese interindividuellen Differenzen kann der Einfluss von unkontrollierter Sekundärvarianz sein.

Bresnitz (1984) unterscheidet den intrinsic und den extrinsic information flow (siehe Abschnitt „Lerntheoretische Grundlagen des Cry Wolf Effektes“). Beim sogenannten „intrinsic information flow“ wird das nicht Eintreten einer negativen Konsequenz nach einer Rückmeldung auf die Ineffektivität des Alarmsystems selbst attribuiert. Die Konsequenz ist, dass der Alarm nicht mehr als effektiv bewertet und daher auf eine tatsächlich folgende Bedrohung nicht mehr angemessen reagiert wird. Dies

entspräche dem Cry Wolf Effekt im eigentlichen Sinne. Beim „extrinsic information flow“ wird die ausbleibende Konsequenz nach einer Rückmeldung nicht auf ein falsch warnendes Alarmsystem zurückgeführt. D.h. die Ursache wird nicht „intrinsic“ innerhalb des Rückmeldesystems gesucht. Stattdessen wird die Situation, vor der das Alarmsystem warnt als nicht gefährlich wahrgenommen, der Grund wird außerhalb des Rückmeldesystems, eben „extrinsic“, gesehen. Hier bleibt das Alarmsystem im eigentlichen Sinne effektiv, jedoch ändert sich für die Versuchspersonen das Konzept der Gefahr. Versuchspersonen reagieren in diesem Fall zwar auf die Rückmeldung, identifizieren die Situation aber als nicht reagierenswert.

Die Art der Situationsgestaltung in Studie 2 begünstigt eine extrinsische Wahrnehmung. So ähnelt die Fehlersituation, in welcher die Automation nicht auf ein bremsendes Vorderfahrzeug reagiert, stark den vier Rückmeldesituationen in der Bedingung mit niedrigem Cry Wolf Potenzial (niedrige Rückmeldungshäufigkeit), welche gleichzeitig auch Teilmenge der 12 Alarmsituationen der Bedingung hohes Cry Potenzial (hohe Rückmeldungshäufigkeit) sind.

Es ist möglich, dass die Versuchspersonen die Situation mit dem Automationsfehler zunächst als nicht gefährlich bewerten, da diese in den ersten Sekunden wie die bereits vorher erlebten Situationen aussieht, in denen die Automation jedoch schließlich bremst. Die Gefahr wird hier zu spät erkannt und entsprechend spät darauf reagiert. Es ist möglich, dass diese Art der Attribution für einen Teil der Versuchspersonen zutraf und daher nicht immer ein Cry Wolf Effekt im eigentlichen Sinne beobachtet wurde. Der Alarm bzw. die Rückmeldung blieb insofern effektiv, dass sie die Aufmerksamkeit zwar zur Fahraufgabe lenkt, die Situation wird dann aber als bekannt und daher nicht gefährlich eingeschätzt.

Der mögliche Einfluss der Sekundärvarianz wird in der Diskussion der weiteren Ergebnisse berücksichtigt.

**Ziel 2:** Untersuchung moderierender Einflüsse der Rückmeldungsart auf die Entstehung eines Cry Wolf Effektes anhand der Entwicklung der Nebenaufgabenleistung.

Als zweite abhängige Variable wird die Entwicklung der Aufmerksamkeitsverteilung zwischen der Fahraufgabe und einer Nebenaufgabe betrachtet. Im Allgemeinen kann festgestellt werden, dass in beiden Rückmeldungsbedingungen die Nebenaufgabenleistung in Rückmeldesituationen signifikant unter der in Situationen ohne Rückmeldung liegt.

Dies spricht zum einen dafür, dass die Nebenaufgabenleistung ein sensibles Maß für die Wirkung von Rückmeldungen ist und zum anderen, dass beide Rückmeldungsarten eine Wirkung auf die Aufmerksamkeitsverteilung haben. Es wurde erwartet, dass sich ein Cry Wolf Effekt in einer von Rückmeldesituation zu Rückmeldesituation zunehmenden Hinwendung zur Nebenaufgabe äußert. In einer Unsicherheitsbedingung wurde entsprechend mit einer geringeren Hinwendung zu einer Nebenaufgabe in Rückmeldesituationen gerechnet als in der Vergleichsbedingung mit normalem Alarm.

Trotz einiger tendenzieller Unterschiede auf deskriptiver Ebene, kann zwischen den beiden Rückmeldungsarten, Unsicherheit und normaler Alarm, keine unterschiedliche Wirkung auf die Nebenaufgabenleistung beobachtet werden. Es wird auch kein signifikanter Einfluss der Rückmeldungshäufigkeit festgestellt, der zumindest für normale Alarmerwartungen erwartet wurde. In der Bedingung niedrige Rückmeldungshäufigkeit steigt die Nebenaufgabenleistung in der Unsicherheitsbedingung nicht über die 4 Rückmeldungssituationen an. In der Bedingung normaler Alarm hingegen hat es zunächst den Anschein, dass es einen systematischen Anstieg gibt. Diese scheinbar erwartungskonforme Interaktion ist aber nicht einmal tendenziell signifikant. In der Bedingung hohe Rückmeldungshäufigkeit steigt die Nebenaufgabenleistung in beiden Rückmeldungsbedingungen gleichförmig über die Fahrt an. Die durchschnittliche Nebenaufgabenleistung liegt in der Bedingung Unsicherheit zwar recht systematisch unter der, der Bedingung normaler Alarm, jedoch kann, unter den gegebenen Bedingungen, kein signifikanter Effekt beobachtet werden. Diese Ergebnisse scheinen zunächst gegen die in Ziel zwei formulierten Hypothesen zu sprechen. Es gibt keinen systematischen Unterschied in Abhängigkeit von der Rückmeldungsart und auch die Unsicherheitsrückmeldung ist, zumindest bei hoher Rückmeldungshäufigkeit, anfällig für einen Cry Wolf Effekt. Ob es sich bei dem Anstieg der Nebenaufgabenleistung aber überhaupt um einen Cry Wolf Effekt handelt ist fraglich. So gibt es keine Interaktion aus dem Kontrollfaktor Situationsklasse und der Wiederholung. D.h. es ist, aufgrund der hier vorliegenden Daten, nicht davon auszugehen, dass der Anstieg der Nebenaufgabenleistung in den Rückmeldungssituationen deutlicher ausfällt als in den Referenzsituationen ohne Rückmeldung. Die Nebenaufgabenleistung liegt zwar in den Rückmeldungssituationen insgesamt unter der Nebenaufgabenleistung der Situationen ohne Rückmeldung, es findet aber ein vergleichbarer Anstieg über die Fahrt statt. Dies deutet darauf hin, dass der beobachtete Anstieg mit Trainingseffekten bzw. einer Optimierung der Nebenaufgabenbearbeitung durch die Versuchspersonen erklärt werden kann.

Der bereits zuvor diskutierte mögliche Einfluss der Sekundärvarianz durch Attribuierung auf die als nicht gefährlich bewertete Situation, kann ebenfalls einen Einfluss haben. Wenn eine Situation als bereits bekannt identifiziert und als ungefährlich eingeschätzt wird, findet mit jedem weiteren Erleben dieser genau gleichen Situation eine weitere Abwendung zu der Nebenaufgabe statt, und zwar unabhängig von der Rückmeldungsart.

In Kombination mit der relativ kleinen Stichprobengröße, können so wohlmöglich vorhandene Effekte, wie der augenscheinliche Unterschied der Nebenaufgabenentwicklung zwischen Unsicherheit und normalem Alarm bei niedriger Rückmeldungshäufigkeit, nicht herausgestellt werden.

**Ziel 3:** Untersuchung moderierender Einflüsse der Rückmeldungsart und der Rückmeldungshäufigkeit auf die Entstehung eines Cry Wolf Effektes, anhand der Betrachtung der Entwicklung des situativen Vertrauens in die Automation.

Als weiteres Maß zur Beurteilung differenzieller Wirkungen von Unsicherheitsrückmeldung und normalem Alarm, wurde das situative Vertrauen in die Automation unmittelbar nach Alarmsituationen erhoben. Es wurde vermutet, dass es aufgrund der Rückmeldung zu einer situativen Vertrauensanpassung kommt, die das Vertrauen in die Automation in Rückmeldungssituationen temporär absinken lässt. Ein Cry Wolf Effekt kann sich dann durch eine Zunahme des situativen Vertrauens in die Automation in den Rückmeldesituationen auswirken, indem die situative Vertrauensanpassung durch die Rückmeldung nicht mehr funktioniert und eine generelle Neigung zum Übervertrauen entsteht.

Wie bereits bei der Betrachtung der Nebenaufgabenbearbeitung, kann ein deutlicher Haupteffekt des Faktors Situationsklasse beobachtet werden. In Rückmeldungssituationen wird der Automation tatsächlich weniger vertraut als in Situationen ohne Rückmeldung. Es kann aber auch ein Interaktionseffekt aus Situationsklasse und Wiederholung festgestellt werden. Das situative Vertrauen in Rückmeldesituationen gleicht sich über die Fahrt dem situativen Vertrauen in Situationen ohne Rückmeldung an. Dies ist unabhängig von der Rückmeldungsart und der Rückmeldungshäufigkeit, da keine Interaktion aus Rückmeldungsart und Wiederholung identifiziert werden kann.

Diese Ergebnisse sprechen insgesamt gegen die postulierten Hypothesen.

Im Gegenteil kann sogar ein Haupteffekt der Rückmeldungsart auf das situative Vertrauen beobachtet werden, wenn beide Situationen mit und ohne Rückmeldung zusammen betrachtet werden. Demnach liegt bei Unsicherheit ein höheres situatives Vertrauen in die Automation vor, als bei einem normalen Alarm. Bei jeweils einzelner Betrachtung der Situationsklasse „Situationen mit Rückmeldung“ ist allerdings kein

signifikanter Unterschied zwischen den Rückmeldungsarten feststellbar. Trotzdem weisen diese Ergebnisse eher in gegenteilige Richtung als in den Hypothesen vermutet.

In der Bedingung mit normalem Alarm und hoher Rückmeldehäufigkeit würde somit eigentlich die stärkste Anpassung des situativen Vertrauens hin zum Niveau in Situationen ohne Rückmeldung erwartet. Es kommt zwar auch zu einem Anstieg des situativen Vertrauens, jedoch fällt dieser augenscheinlich deutlich geringer aus als in beiden Unsicherheitsbedingungen.

Dies ist ein möglicher Hinweis auf eine alternative Hypothese, dass die Wahrnehmung einer Rückmeldung als ineffektiv, sich dämpfend auf das situative Vertrauen in Rückmeldesituationen auswirkt. D.h. werden bei einem normalen Alarm zu viele Rückmeldungen ohne Automationsfehler erlebt, so steigt das situative Vertrauen weniger als in den vergleichbaren Situationen mit Unsicherheitsrückmeldung.

Auch bei der Betrachtung des situativen Vertrauens kann der Einfluss von Sekundärvarianz, durch die Darbietung immer gleicher Fahrsituationen, einen konfundierenden Einfluss haben. Der Anstieg des situativen Vertrauens in die Automation könnte so z.B. darauf zurückzuführen sein, dass die Automation als so leistungsfähig eingeschätzt wird, dass sie, trotz Rückmeldung eines kritischen Zustandes, die Situation bewältigen kann. In diesem Fall ändert sich das Konzept der Gefahr und weniger die Einstellung gegenüber dem Rückmeldesystem.

Auffällig ist der starke Anstieg des situativen Vertrauens von der ersten Rückmeldesituation auf die nächsten Rückmeldesituationen. Obwohl die Versuchspersonen noch gar nicht die Erfahrung machen konnten, dass die Automation in einer bestimmten Situation trotz Rückmeldung immer richtig reagiert, wächst das situative Vertrauen am stärksten von der ersten zur zweiten Situation mit Rückmeldung.

Eine mögliche Erklärung ist, dass es sich bei diesem Anstieg weniger um einen Cry Wolf Effekt im eigentlichen Sinne handelt, sondern, dass Versuchspersonen aufgrund der Instruktion mit deutlich gefährlicheren Situationen im Fall einer Automationsrückmeldung gerechnet haben, bzw. hier eine große Ungewissheit vorlag, wie eine solche Situation wohl aussehen könne. Diese Erklärung wäre konform mit dem Erklärungsansatz der Sekundärvarianz und der Veränderung des Konzeptes der Gefahr im Kontext des teil-/hochautomatisierten Fahrens. Um hierzu aber konkretere Aussagen treffen zu können, muss in Folgestudien der vermutete Einfluss von Sekundärvarianz kontrolliert werden.

**Ziel 4:** Untersuchung der Auswirkungen von Rückmeldungsart und Rückmeldungs-häufigkeit auf das Gesamtvertrauen in die Automation.

Bei der Betrachtung des Gesamtvertrauens nach der Fahrt, wurde eigentlich erwartet, dass sich eine Unsicherheitsrückmeldung, verglichen mit einem normalen Alarm, positiv auf die Bewertung auswirkt. Der Grund wurde darin vermutet, dass ein normaler Alarm aufgrund des Cry Wolf Effektes als unzuverlässig erlebt wird und zu einer schlechteren Kontrollierbarkeit in der Situation mit Automationsfehler führt. Da aber insbesondere in den Ergebnissen zur Kontrollierbarkeit kein wirklich deutlicher Unterschied zwischen den Rückmeldungsarten herausgestellt wurde, konnte wohlmöglich in keiner der erfassten Vertrauensdimensionen ein systematischer Unterschied zwischen den Rückmeldungsarten beobachtet werden. Auch konnte kein systematischer Unterschied aufgrund der Rückmeldungshäufigkeit identifiziert werden.

Insgesamt fällt auf, dass das Gesamtvertrauen in die Automation sehr hoch ist, obwohl kurz vor dem Ausfüllen des Fragebogens eine Situation erlebt wurde, in der ein Automationsfehler eher schlecht kontrolliert werden konnte.

Grund für das hohe Gesamtvertrauen könnte die Begünstigung der Entstehung von Übervertrauen in die Automation, durch wiederholtes Erleben von diversen gleichartigen Rückmeldungssituationen ohne Automationsfehler, sein.

Da diese Situationen wiederholt von der Automation gemeistert wurden, wurde nicht die Rückmeldung als ineffizient angesehen, sondern die Automation als besonders sicher. Dieses Übervertrauen kann anscheinend auch durch das Erleben einer Situation mit Automationsfehler nicht besonders negativ beeinflusst werden.

**Ziel 5:** Untersuchung der Auswirkungen von Rückmeldungsart und Rückmeldungshäufigkeit auf die Akzeptanz der Automation.

Die als Akzeptanzmaß erhobene Nutzungstendenz der Automation konnte keinen Unterschied zwischen den beiden Rückmeldungsbedingungen und auch keinen Einfluss der erlebten Rückmeldungshäufigkeit zeigen. Die Versuchspersonen tendieren dazu, unabhängig von Rückmeldungsart und Häufigkeit, die Automation eher eingeschaltet zu lassen. Diese Indifferenz widerspiegelt die Ergebnisse des Gesamtvertrauens, aber auch die eher geringen Unterschiede in den übrigen abhängigen Variablen.

**Zusammengefasst** kann in Studie 2 kein eindeutiger Beleg dafür gefunden werden, dass die Unsicherheitsrückmeldung anders interpretiert bzw. verarbeitet wird als ein normaler Alarm. Deskriptiv und teilweise statistisch abgesichert, weisen zwar einige Ergebnisse in Richtung der formulierten Hypothesen, dennoch lässt sich bei vielen Maßen nicht der vermutete Effekt beobachten. Werden auf der einen Seite Vorteile aufgrund der Unsicherheitsrückmeldung beobachtet, so werden diese häufig an



anderer Stelle durch keine beobachtbaren Unterschiede relativiert, oder positive Effekte sind so gering, dass daraus kein nennenswerter Vorteil bzw. Sicherheitsgewinn zu verbuchen ist. Zumindest führt die Unsicherheitsrückmeldung nicht zu schlechteren Ergebnissen als ein normaler Alarm.

Die deskriptiv beobachteten Unterschiede zwischen Unsicherheitsrückmeldung und dem normalen Alarm scheinen aber systematischer Art zu sein, auch wenn diese nicht immer durch Tests abgesichert werden können. D.h. wenn es einen zumindest deskriptiven Unterschied gibt, dann besteht dieser aus einem tendenziellen Vorteil gegenüber dem normalen Alarm. Die Gründe für die wenigen abgesicherten Unterschiede liegen mit großer Wahrscheinlichkeit in methodischen Problemen, deren Bedeutung aber erst mit der Auswertung der Daten offensichtlich wurde.

Als problematisch auf die Auswertung der Studie wirkt sich die Verringerung der Teilstichprobengröße durch nachträgliche Aussonderung von Versuchspersonen aus. Die Power der angewendeten Analyseverfahren ist dadurch eingeschränkt und mögliche Unterschiede können somit nicht eindeutig herausgestellt werden. Die daraus gezogene Konsequenz ist, in weiteren Studien bereits während der Versuchsdurchführung, Versuchspersonen mit einer Tendenz zur „vorschnellen“ Reaktion zu identifizieren und dann zusätzliche Versuchspersonen in die Gruppe aufzunehmen, bis je Bedingung eine ausreichende Gruppenstärke von Personen erreicht ist.

Ein zweiter wesentlicher Aspekt ist der mögliche Einfluss der Sekundärvarianz durch die Gleichartigkeit der Fahrsituationen. In Folgestudien muss diese Art der Sekundärvarianz berücksichtigt und entsprechend kontrolliert werden. Eine höhere Variabilität der erlebten Situationen als in Studie 2 umgesetzt, ist zudem deutlich realitätsnäher. So wird es im Verlauf einer realen Fahrt kaum möglich sein, dass mehrere aufeinander folgende Situationen wirklich exakt gleich sind.

## **9 Studie 3**

### **9.1 *Ziele Studie 3***

Studie 3 hat mehrere Ziele und wird daher so aufgebaut, dass mehrere unterschiedliche Fragestellungen adressiert werden können. Um eine Übersicht über den Aufbau der Studie zu bekommen, werden folgend die Ziele der Studie und die damit verbundenen Forschungsfragen und Hypothesen aufgeführt.

**Ziel 1:** Wiederholte Betrachtung des Vergleichs der Alarmstrategien Unsicherheitsrückmeldung und normaler Alarm hinsichtlich ihrer Auswirkung auf die Entstehung eines Cry Wolf Effektes.

Bei dieser wiederholten Betrachtung soll die in Studie 2 vermutete Sekundärvarianz, durch die wiederkehrend gleichartige Gestaltung der Alarmsituation, besser kontrolliert werden. (siehe Diskussion Studie 2).

**Forschungsfrage:** Die Forschungsfrage ist die gleiche wie bereits in Studie 2. Unterscheidet sich eine Unsicherheitsrückmeldung überhaupt von einem normalen Alarm hinsichtlich der Anfälligkeit für einen Cry Wolf Effekt?

**Hypothese:** Die Hypothese ist, aufgrund der Wiederholung, die gleiche wie in Studie 2: Wenn die Unsicherheitsrückmeldung anders interpretiert wird als ein normaler Alarm ist die Neigung zu einem Cry Wolf Effekt bei gleich vielen erlebten Rückmeldungen ohne Automationsfehler bei einer Unsicherheitsrückmeldung geringer als bei einem normalen Alarm. Das heißt, wenn über die Fahrt nur Rückmeldungen ohne nachfolgenden Automationsfehler erlebt werden und in der letzten Fahrsituation tatsächlich ein Automationsfehler nach einer Rückmeldung folgt, dann ist bei Unsicherheitsrückmeldung die Erwartung eines Automationsfehlers höher und die Kontrollierbarkeit besser als in einer Bedingung mit normalem Alarm. Es wird dann häufiger und/oder besser auf einen Automationsfehler reagiert als in einer Bedingung normaler Alarm.

**Ziel 2:** Bestimmung der absoluten Größe eines Cry Wolf Effektes bei Unsicherheitsrückmeldung und normalem Alarm durch einen Vergleich mit einem idealen Alarm.

Bisher konnte nur die relative Größe eines Cry Wolf Effektes durch den Vergleich zwischen der Unsicherheitsrückmeldung und dem normalen Alarm ermittelt werden. Es konnte aber keine Aussage getroffen werden, wie weit beide Rückmeldungsarten von einem theoretischen Optimum, nämlich gar keinem Cry Wolf Effekt, entfernt sind. Beide Rückmeldungsbedingungen werden nun mit einer Rückmeldebedingung verglichen, in der kein Cry Wolf Effekt entstehen kann. Diese Bedingung wird durch einen idealen Alarm umgesetzt, bei dem nur im Fall eines tatsächlichen Automationsfehlers eine Rückmeldung gegeben wird. Dies ermöglicht eine Aussage darüber, wie groß der Cry Wolf Effekt auch in der Bedingung Unsicherheitsrückmeldung insgesamt ist.

**Forschungsfrage:** Gibt es in der Bedingung Unsicherheitsrückmeldung überhaupt einen Cry Wolf Effekt und wie groß ist ein möglicher Cry Wolf Effekt bei Unsicherheitsrückmeldung absolut?

**Hypothese:** Wenn die Unsicherheitsrückmeldung robust gegen die Entstehung eines Cry Wolf Effektes ist, dann ist die Reaktion auf einen Automationsfehler in einer Bedingung mit Unsicherheitsrückmeldung nicht

signifikant schlechter als in einer Bedingung mit idealem Alarm. Gemäß der Hypothese aus Ziel 1, ist die Reaktion auf den Automationsfehler in der Bedingung normaler Alarm schlechter als in den Bedingungen idealer Alarm und Unsicherheitsrückmeldung.

**Ziel 3:** Untersuchung rückmeldungsartspezifischer Lerneffekte aufgrund von Evidenzevaluation.

Die Formulierung einer Rückmeldung als Unsicherheit erzeugt keine eindeutige Zusammenhangshypothese zwischen Rückmeldung und erwarteter Konsequenz (Automationsfehler). Ein normaler Alarm lässt eine solche Zusammenhangshypothese jedoch annehmen. Ausbleibende Automationsfehler stellen entsprechend negative Evidenz für diese Ausgangshypothese dar, bis diese so schwach ist, dass die Rückmeldung als nutzlos erachtet wird (ein Cry Wolf Effekt entsteht). Wird jedoch positive Evidenz erlebt, so wirkt diese der bereits erfahrenen negativen Evidenz entgegen.

**Forschungsfrage:** Wirkt sich die Erfahrung einer Situation mit tatsächlichem Automationsfehler (korrekter Alarm bzw. positive Evidenz) unterschiedlich auf die Entstehung weiterer Cry Wolf Effekte bei den Rückmeldungsarten Unsicherheitsrückmeldung und normaler Alarm aus?

**Hypothese:** Ein normaler Alarm erzeugt eine initiale gerichtete Zusammenhangshypothese zwischen Rückmeldung und dem Eintreten eines Automationsfehlers. Das Erleben von Rückmeldungen ohne Automationsfehler stellt demnach negative Evidenz für diese initiale Zusammenhangshypothese dar. Bei wiederholtem Erleben negativer Evidenz ändert sich die Zusammenhangshypothese – ein Cry Wolf Effekt entsteht. Wird jedoch unerwartet positive Evidenz erlebt (Automationsfehler nach Rückmeldung), so wird diese, aufgrund des aversiven Charakters, stark gewichtet und im weiteren Verlauf der Fahrt besteht große Unsicherheit bezüglich des Zusammenhangs zwischen Rückmeldung und Automationsfehler. Durch die große Unsicherheit wird das Bilden einer neuen Zusammenhangsregel gehemmt, der Cry Wolf Effekt wird minimiert. Dies äußert sich in:

- Einer deutlich besseren Reaktion in einer zweiten Automationsfehlersituation im weiteren Fahrtverlauf nach Erleben eines Automationsfehlers
- Weniger bearbeiteten Nebenaufgaben im weiteren Fahrtverlauf nach Erleben eines Automationsfehlers als vor dem Automationsfehler bearbeitet wurden

- Geringeres situatives Vertrauen in die Automation in Rückmeldesituationen im weiteren Fahrtverlauf nach Erleben eines Automationsfehlers

Bei der Unsicherheitsrückmeldung hingegen liegt, aufgrund der Rückmeldungsgestaltung, von vornherein eine große Unsicherheit bezüglich des Zusammenhangs zwischen Rückmeldung und Automationsfehler vor. Daher gibt es keinen Unterschied in Reaktionszeit, Nebenaufgaben und situativen Vertrauen vor und nach dem Erleben eines ersten Automationsfehlers. Oder, wenn ein Unterschied vorliegt, ist dieser deutlich geringer als in der Bedingung normaler Alarm.

**Ziel 4:** Untersuchung der Effektivität einer gemischten Alarmstrategie aus Unsicherheitsrückmeldung und normalem Alarm.

Es wird untersucht, ob sich dreistufige Rückmeldungen aus den Stufen keine Rückmeldung, Wahrscheinlichkeit bzw. Unsicherheit und Alarmstufe, wie sie auch Likelihoodalarme darstellen, negativ auf die Effektivität der Warnstufe bzw. Unsicherheitsstufe auswirken. Begründet wird diese Vermutung damit, dass in bisherigen Studien gezeigt werden konnte, dass bei Likelihoodalarmen die Effektivität der Alarmstufe auf Kosten der Effektivität der Warnstufe gesteigert werden konnte (siehe „Theoretischer Hintergrund“). Es soll somit der Frage nachgegangen werden, ob für Situationen, in denen die Automation sicher ist, dass sie die Kontrolle nicht angemessen ausführen kann, anstelle der Unsicherheitsrückmeldung ein dringlicher Alarm gegeben werden sollte, oder ob diese Einteilung entsprechend zu einer Verringerung der Effektivität der Unsicherheitsrückmeldung führen würde.

**Forschungsfrage:** Wenn in einer ersten Situation mit Automationsfehler mit einem normalen Alarm gewarnt wird, wirkt sich dies negativ auf die Effektivität einer Unsicherheitsrückmeldung aus, wenn in einer zweiten Situation mit Automationsfehler Unsicherheit zurückgemeldet wird.

**Hypothese:** Das Mischen unterschiedlicher Rückmeldungsarten, lässt eine Erwartungshypothese im Sinne eines Grundes für die Unterscheidung zweier Rückmeldungsarten entstehen. Wenn in „False Alarm“ Situationen im Regelfall Unsicherheit zurückgemeldet wird und dann in einer korrekten Alarmsituation ein normaler Alarm, reduziert dies die Effektivität der Unsicherheitsrückmeldung, wenn diese in einer folgenden korrekten Alarmsituation zurückgemeldet wird.

**Ziel 5:** Untersuchung der Auswirkung einer Unsicherheitsrückmeldung auf die Erwartung eines Automationsfehlers in Situationen ohne Rückmeldung (Reliance). durch Betrachtung der Effektivität einer Unsicherheitsrückmeldung in einer „Miss-Situation“.

Eine Miss-Situation ist eine Situation, in der trotz Automationsfehler kein Alarm geäußert wird.

Bei einem Alarmsystem unterscheidet z.B. Meyer (2001) zwei unterschiedliche Formen der Einstellung gegenüber dem Alarmsystem, die Compliance und die Reliance. Reliance beschreibt dabei den Grad der Erwartung, dass das Alarmsystem in einer Situation mit Automationsfehler auch tatsächlich warnt. Die Compliance bezieht sich auf den Grad der Erwartung, dass wenn ein Alarm ertönt, es in diesem Fall auch zu einem Automationsfehler kommt. Ein Cry Wolf Effekt bezieht sich folglich auf die Compliance. Ist die Reliance in das Alarmsystem hoch, sollte entsprechend auch die Reliance in die Fahrzeugautomation hoch sein, da in Situationen, in denen kein Alarm geäußert wird, nicht mit einem Automationsfehler gerechnet wird. In der Literatur finden sich Hinweise, die eine gewisse Abhängigkeit der Compliance und Reliance vermuten lassen (z.B. Dixon, Wickens & McCarly, 2006 & 2007). Da der Cry Wolf Effekt ein Fall besonders niedriger Compliance ist (teilweise sogar eine gegenteilige Überzeugung, dass wenn ein Alarm ertönt, es NICHT zu einem Automationsfehler kommt) und postuliert wurde, dass der Cry Wolf Effekt von der Rückmeldungsart abhängt, kann es sein, dass die Reliance somit indirekt auch durch die Art der Rückmeldung beeinflusst wird (unter Annahme, dass Compliance und Reliance tatsächlich voneinander abhängen).

**Forschungsfrage:** Wirkt sich die Art der Rückmeldungsgestaltung durch unterschiedlich ausgeprägte Anfälligkeit für einen Cry Wolf Effekt und dadurch unterschiedlich hohe Compliance auch unterschiedlich auf die Höhe der Reliance in das Alarmsystem aus?

**Hypothese:** Das Erleben eines ausgeprägten Cry Wolf Effektes führt in nachfolgenden Bedingungen zu einem generell niedrigeren Vertrauen in das Alarmsystem, d.h. auch in einer Situation ohne Alarm wird ein Automationsfehler stärker erwartet und die Kontrollreaktion ist besser als bei einer Alarmstrategie in der vorher kein so großer Cry Wolf Effekt erlebt wurde. Die Unsicherheitsrückmeldung müsste entsprechend in Situationen mit Automationsfehler ohne Rückmeldung zu einer schlechteren Kontrollierbarkeit führen als in einer Bedingung mit normalem Alarm. Entsprechend müsste die Kontrollierbarkeit einer solchen Situation in einer Bedingung idealer Alarm noch schlechter sein.

## 9.2 Methode Studie 3

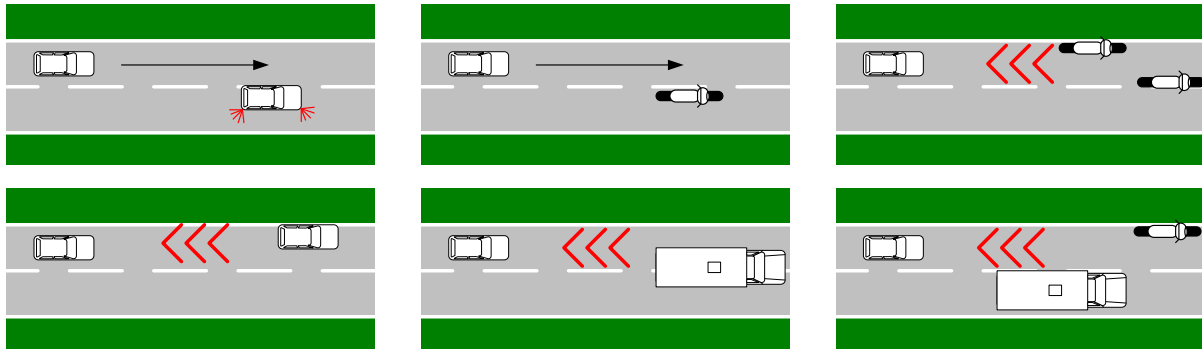
Um die unterschiedlichen Forschungsfragen und daraus abgeleitete Hypothesen in einer Studie adressieren zu können, wird ein Gesamtstudiendesign gewählt, das die jeweils einzelne Betrachtung der Forschungsfragen bzw. Hypothesen ermöglichen soll. Es werden daher unterschiedliche Versuchsdesigns aufeinander aufgebaut und in einem übergeordneten Studiendesign vereinigt.

Die Studie wird, wie bereits auch die Studien 1 und 2, in einem fixed base Fahrsimulator durchgeführt (siehe Gesamtmethodenteil). Für die Umsetzung der in Studie 3 vorgenommenen Simulatorfahrten, wird auf das im übergeordneten Methodenteil beschriebene Grunddesign zurückgegriffen. Das heißt, Versuchspersonen fahren in einem Fahrzeug mit einer integrierten Längs- und Querverführung über eine nebelige zweispurige und gerade Autobahn und erleben während der Fahrt klare und unklare Situationen mit einem Fremdfahrzeug.

### 9.2.1 Kontrolle von Sekundärvarianz

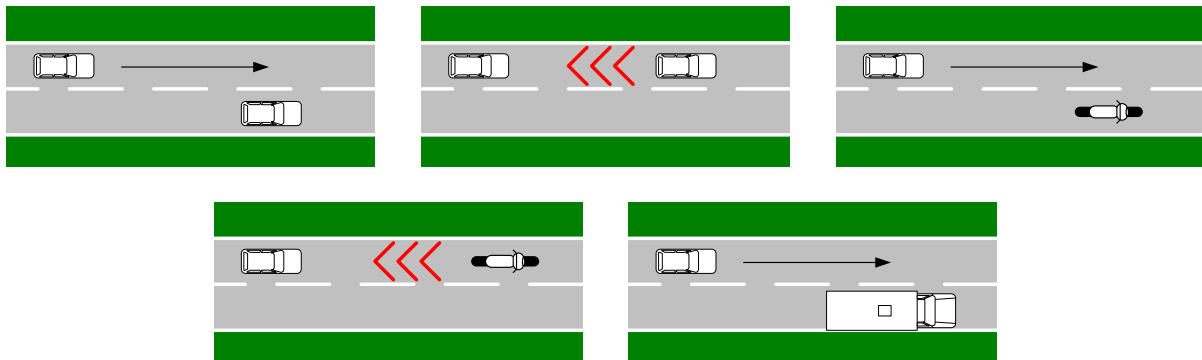
#### 9.2.1.1 Fahrsituationen

In Studie 2 wurde ein systematischer Einfluss von Sekundärvarianz durch die Gleichartigkeit der unklaren Situationen ohne Automationsfehler und der Situation mit Automationsfehler vermutet. Es wurde vermutet, dass durch das wiederkehrende Erleben stets der gleichen Situation, eine Attribution auf die Situation vorgenommen wird und dadurch Versuchspersonen diese spezielle Situation als nicht mehr gefährlich wahrnehmen. Zur Minimierung der Sekundärvarianz wird daher eine höhere Variabilität der Fahrsituationen insgesamt und der unklaren Situationen insbesondere geschaffen. Abbildung 43 zeigt die in Studie 3 insgesamt verwendeten Rückmeldungssituationen mit Rückmeldung ohne Automationsfehler.



**Abbildung 43: Unklare Situationen ohne Automationsfehler in Studie 3: Oben links: versetzt fahrendes Fahrzeug auf rechtem Fahrstreifen, Automation fährt berechtigt daran vorbei (FA1); Oben Mitte: versetzt fahrendes Motorrad auf dem rechten Fahrstreifen, Automation fährt berechtigt daran vorbei (FA2); Oben rechts: Zwei versetzt fahrende Motorräder auf dem eigenen Fahrstreifen, Automation fährt berechtigt in Folgefahrt (FA3); Unten links: versetzt fahrendes Fahrzeug auf dem eigenen Fahrstreifen, Automation fährt berechtigt in Folgefahrt (FA4); unten Mitte: versetzt fahrender LKW auf eigenem und Nachbarfahrstreifen, Automation fährt berechtigt in Folgefahrt (FA5); Unten rechts: Versetzt fahrender LKW auf rechtem Fahrstreifen und versetzt fahrendes Motorrad auf eigenem Fahrstreifen, Automation fährt berechtigt in Folgefahrt (FA6)**

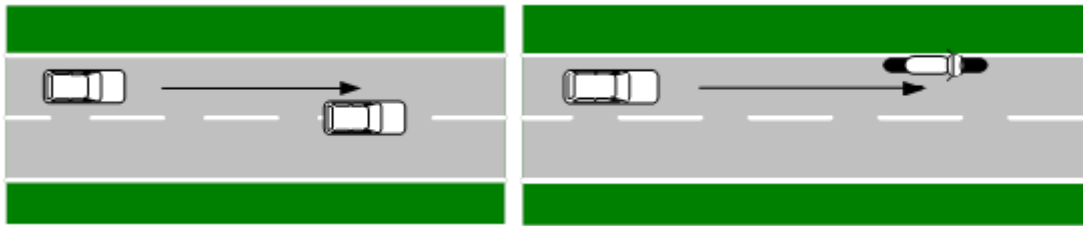
Auch die Anzahl der zur Verfügung stehenden klaren Situationen wird erhöht, so dass in der Fahrt möglichst eine Vielzahl unterschiedlicher Situationen vorkommt, und ein und dieselbe Situation nicht zweimal direkt nacheinander erlebt wird (Abbildung 44).



**Abbildung 44: Klare Situationen in Studie 3: Oben links: Fahrzeug auf rechtem Fahrstreifen, Automation fährt daran vorbei (N1); Oben Mitte: Fahrzeug auf eigenem Fahrstreifen, Automation geht in Folgefahrt (N2); Oben rechts: Motorrad auf rechtem Fahrstreifen, Automation fährt daran vorbei (N3); Unten links: Motorrad auf eigenem Fahrstreifen, Automation geht in Folgefahrt (N4); Unten rechts: LKW auf rechtem Fahrstreifen, Automation fährt daran vorbei (N5).**

Als unklare Situation mit Automationsfehler kann aus zwei verschiedenen Varianten gewählt werden. In beiden Situationen müssen Versuchspersonen mit einer Bremsung reagieren, um eine Kollision zu vermeiden (Abbildung 45). Beide

Situationen weisen, durch gleiche Geschwindigkeiten und TTC zum Zeitpunkt des Erscheinens der Vorderfahrzeuge, die gleiche Kritikalität auf.

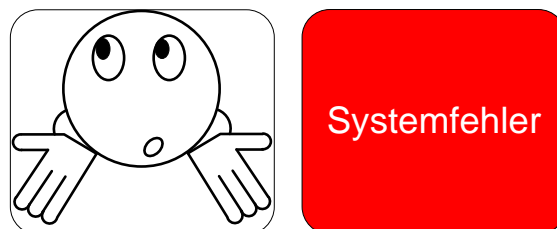


**Abbildung 45: Varianten der unklaren Situation mit Automationsfehler: Ein Fahrzeug fährt deutlich versetzt auf linkem und rechtem Fahrstreifen zugleich, die korrekte Automationsaktion hier wäre in die Folgefahrt zu gehen, in dieser Situation reagiert die Automation jedoch nicht auf das Vorderfahrzeug.**

Während der Fahrt sollen die Versuchspersonen wieder eine visuelle Nebenaufgabe bearbeiten, um über die Nebenaufgabenleistung ein Maß für die Aufmerksamkeitsverteilung zwischen Fahr- und Nebenaufgabe zu erlangen. Anreiz zur Bearbeitung möglichst vieler Nebenaufgaben, ohne dabei die Fahraufgabe zu vernachlässigen, ist wieder das im Gesamtmethodenteil beschriebene Belohnungs-Bestrafungs-Paradigma. Die Versuchspersonen werden nach jeder Situation mit einem Fremdfahrzeug befragt, wie sehr sie der Automation in dieser Situation vertraut haben (siehe Gesamtmethodenteil).

#### **9.2.1.2 Alarmverständnis**

Um Sekundärvarianz durch variierendes Verständnis der Rückmeldungssymbole zu eliminieren, wird, wie bereits in Studie 1 und 2, die Bedeutung der jeweiligen Symbole in der Instruktion erklärt. D.h. bei dem Unsicherheitssymbol werden die Versuchspersonen informiert, dass die Automation unsicher bezüglich ihrer situativen Leistungsfähigkeit sei. Bei einem normalen Alarm Symbol wird den Versuchspersonen gesagt, dass in dieser Situation ein Systemfehler resultieren kann.



**Abbildung 46: Links: Unsicherheitssymbol; rechts: Symbol eines normalen Alarms.**



## 9.2.2 Versuchsdesign je Teilziel

### 9.2.2.1 Ziele 1, 2 & 4-Teil 1: Design

#### **Vergleich von Unsicherheitsrückmeldung, normalem Alarm, idealem Alarm und Mischstrategie-Teil1:**

Es wird eine Strategie der Unsicherheitsrückmeldung mit einem normalen Alarm, einem idealen Alarm und einer Mischstrategie aus Unsicherheitsrückmeldung und normalem Alarm bezüglich der Anfälligkeit gegenüber einem Cry Wolf Effekt verglichen. Dieser Vergleich bildet den Zwischengruppenfaktor

„Rückmeldungsarten“. Alle Rückmeldungsstrategien werden der gleichen Behandlung, einer Cry Wolf Konditionierung, ausgesetzt. In der, in mehreren Situationen während der Fahrt, eine der Alarmstrategie entsprechende Rückmeldung gegeben wird, auf die kein Automationsfehler folgt. In der Bedingung idealer Alarm wird entsprechend keine Rückmeldung gegeben.

Die Behandlung wird gleichzeitig als Innergruppenfaktor „Cry Wolf Entwicklung“ betrachtet, in dem über die verschiedenen Situationen mit Wiederholungsmessung die Entwicklung eines Cry Wolf Effektes beobachtet wird.

In einer letzten Situation mit Rückmeldung passiert dann tatsächlich ein Automationsfehler. Diese Situation wird als Nachhermessung nach der Behandlung „Cry Wolf Konditionierung“ behandelt. Daraus resultiert ein zweifaktorielles Mixed Models Versuchsdesign mit Messungen auf den einzelnen Stufen des Wiederholungsfaktors und Nachhermessung (Tabelle 29).

**Tabelle 29: Versuchsplan Teilexperiment „Ausmaß eines Cry Wolf Effektes in Abhängigkeit unterschiedlicher Rückmeldungsarten“**

Zwischengruppen- faktor: Rückmeldungsart		Behandlung: Erzeugung eines Cry Wolf Potenzials mit Wiederholungsfaktor (Nebenaufgabenleistung / situatives Vertrauen)					Nachhermessung Kontrollierbarkeit: (Reaktionszeit & Häufigkeit) in Sit 31
		Sit 1	Sit2	Sit3	Sit4	Sit5	
	Unsicherheit	1.1	1.2	1.3	1.4	1.5	1
	Normaler Alarm	2.1	2.2	2.3	2.4	2.5	2
	Idealer Alarm	3.1	3.2	3.3	3.4	3.5	3
	Mischstrategie	4.1	4.2	4.3	4.4	4.5	4

#### 9.2.2.1.1 Unabhängige Variablen

##### **Zwischengruppenfaktor Rückmeldungsart**

Der Faktor „Rückmeldungsart“ hat folgende Stufen:

1. Unsicherheitsrückmeldung: in allen unklaren Situationen wird Automationsunsicherheit zurückgemeldet (Abbildung 46 links), unabhängig davon ob ein Automationsfehler folgt oder nicht.
2. Normaler Alarm: In allen unklaren Situationen wird ein normaler Alarm zurückgemeldet (Abbildung 46 rechts), unabhängig davon ob ein Automationsfehler folgt oder nicht.
3. Idealer Alarm: Nur in Situationen mit einem Automationsfehler wird ein normaler Alarm zurückgemeldet (Abbildung 46 rechts), in allen anderen unklaren Situationen erfolgt keine Rückmeldung.
4. Mischstrategie: In unklaren Situationen ohne Automationsfehler wird zunächst Unsicherheit zurückgemeldet (Abbildung 46 links), in Situationen mit Automationsfehler wird ein normaler Alarm zurückgemeldet (Abbildung 46 rechts) – diese Strategie bildet einen Likelihoodalarm, wie in Abschnitt Likelihoodalarme, Alarmcoding & Task-Critical Information beschrieben, dar.

##### **Wiederholungsfaktor – Cry Wolf Konditionierung**

Die Phase der Cry Wolf Konditionierung ist eine ununterbrochene Fahrt mit 30 Fremdfahrzeugsituationen plus eine 31. Situation als Nachhermessung. In diese 30 Situationen werden 5 Situationen mit Rückmeldungen eingestreut, in denen aber kein Automationsfehler passiert. In der Bedingung idealer Alarm wird in diesen Situationen nichts zurückgemeldet.

Diese 5 Rückmeldungssituationen bilden die Messpunkte des Wiederholungsfaktors – Cry Wolf Konditionierung.

Um ein ausreichend hohes Cry Wolf Potenzial zu erzeugen, wird ein Verhältnis von 1:5 unklaren zu klaren Situationen gewählt. Auf 5 klare Situationen ohne Rückmeldung kommt also eine unklare Situation mit Rückmeldung und ohne Automationsfehler.

Das gewählte Verhältnis aus unklaren und klaren Situationen soll in etwa die Mitte der Verhältnisse der beiden unterschiedlich hohen Cry Wolf Potenziale (Rückmeldungshäufigkeiten) aus Studie 2 treffen. In Studie 2 wurde in der Bedingung hohe Rückmeldungshäufigkeit (hohes Cry Wolf Potenzial) ein Verhältnis von 1:2,33 (12 unklare Situationen auf 28 klare Situationen) gewählt, und in der Bedingung niedrige Rückmeldungshäufigkeit (niedriges Cry Wolf Potenzial) ein Verhältnis aus 1:8,5 (4 unklare Situationen auf 36 klare Situationen).

Die Situationsabfolge wurde für die vier Rückmeldungsbedingungen konstant gehalten (Abbildung 47).

Sit.-Nr.		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		
Sit.-Art		N1	N5	N3	FA 2	N2	N4	N1	N5	N3	FA 5	N5	N2	N1	N1	N2	N1	N5	FA 1	N1	N1	N4	N2	FA 3	N5	N1	N3	FA 4	N1	N4	N1	CA 1		
Faktorstufen Rückmeldungsart	Unsicherheits- rückmeldung	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	○	○	●	
	Normaler Alarm	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	●	
	Idealer Alarm	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●
	Mischstrategie	○	○	○	●	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	○	○	○	○	●
		○ = Keine Rückmeldung    ● = Unsicherheitsrückmeldung    ● = Normaler Alarm																																

**Abbildung 47: Abfolge von unklaren und klaren Situationen für die verschiedenen Rückmeldungsarten. In den ersten 30 Situationen folgt nach unklaren Situationen kein Automationsfehler, in der 31. Situation folgt ein Automationsfehler.**

### 9.2.2.2 Ziele 3 und 4-Teil 2: Design

Sowohl zur Überprüfung der Hypothesen der Wechselwirkung aus Rückmeldungsart und Erfahrung eines Cry Wolf Effektes auf die weitere Entwicklung eines Cry Wolf Effektes (Evidenzevaluation - Ziel 3), als auch zur Überprüfung der Effektivität einer Rückmeldungs-Mischstrategie (Ziel 4), ist eine Längsschnittbetrachtung unterschiedlicher Zeitpunkte einer Fahrt notwendig. Es wurde daher auf dem Versuchsdesign für die Ziele 1, 2 & 4 aufgebaut und die hier durchgeführte Fahrt (Abbildung 47) um einen weiteren, gleich langen Fahrtabschnitt ergänzt. Es wurde vermutet, dass in Abhängigkeit von der Rückmeldungsart, während der 1. Fahrt unterschiedlich hohe Cry Wolf Effekte erzeugt werden, welche in der Situation mit Automationsfehler zu einer entsprechend unterschiedlichen Erwartung des Automationsfehlers führen. Die Auswirkungen eines möglichen Cry Wolf Effektes können somit von den Versuchspersonen direkt erfahren werden. Diese, von der Rückmeldungsart abhängige, Erfahrung kann die Entwicklung eines Cry Wolf Effektes im weiteren Verlauf der Fahrt maßgeblich beeinflussen.

Es wurde ein zweifaktorielles Mixed Models Versuchsdesign mit der „Rückmeldungsart“ als Zwischengruppenfaktor und der Entwicklung über die verschiedenen Situationen bzw. Zeitpunkte der Fahrt als Innergruppenfaktor „Evidenzerfahrung“ gewählt. Wesentlicher Zweck des Innergruppenfaktors war vor allem der Vergleich zwischen der ersten Kontrollierbarkeitssituation beim ersten Automationsfehler und der zweiten Kontrollierbarkeitssituation beim zweiten Automationsfehler. Des Weiteren sollte durch den Innergruppenfaktor die Erfassung der Entwicklung des situativen Vertrauens und der Nebenaufgabenleistung im ersten Teil der Fahrt, bis zum ersten Automationsfehler, mit der Entwicklung dieser Variablen im zweiten Teil der Fahrt, vom ersten Automationsfehler bis zum zweiten Automationsfehler, verglichen werden können (Tabelle 30).

**Tabelle 30: Versuchsdesign Teilerperiment „Auswirkung positiver Evidenz“**

		Innergruppenfaktor „Evidenzerfahrung“			
		Ohne Evidenzerfahrung (1.Fahrt)		Mit Evidenzerfahrung (2.Fahrt)	
		Behandlung 1 Erzeugung eines Cry Wolf Potenzials	Messung 1 Abhängige Variablen Kontrollierbarkeit / Nebenaufgaben- leistung / situatives Vertrauen	Behandlung 2 Erzeugung eines Cry Wolf Potenzials	Messung 2 Abhängige Variablen Kontrollierbarkeit / Nebenaufgaben- leistung / situatives Vertrauen
Zwischengruppen-faktor: Rückmeldungsart	Unsicherheit	Sit 1 - 30	y1.1	Sit 32 - 61	y1.2
	Normaler Alarm	Sit 1 - 30	y2.1	Sit 32 - 61	y2.2
	Idealer Alarm	Sit 1 - 30	y3.1	Sit 32 - 61	y3.2
	Misch- strategie	Sit 1- 30	y4.1	Sit 32 -61	y4.2

#### 9.2.2.2.1 Unabhängige Variablen

- Zwischengruppenfaktor Rückmeldungsart  
Siehe Abschnitt „Versuchsdesign für Ziele 1, 2 & 4“
- Wiederholungsfaktor Evidenzerfahrung

Der Wiederholungsfaktor „Evidenzerfahrung“ hat zwei Stufen, welche von allen Rückmeldungsartgruppen nacheinander in der gleichen Reihenfolge (erst Stufe 1, dann Stufe 2) erlebt wurden:

Der erste Teil der Fahrt bis einschließlich des ersten Automationsfehlers, stellte die Stufe 1, „Cry Wolf Effekt & Entwicklung ohne Evidenzerfahrung“ dar. Bis zum ersten Automationsfehler wurden in dieser Fahrt nur Rückmeldungen ohne Automationsfehler erlebt. Zur Umsetzung der zweiten Stufe „Cry Wolf Effekt & Entwicklung mit Evidenzerfahrung“, wurde die Fahrt nach dem Erleben des ersten Automationsfehlers fortgesetzt. In diesem zweiten Teil der Fahrt wurden wieder insgesamt 31 klare und unklare Fahrsituationen durchfahren, in den ersten 30 Situationen fanden wieder 5 Rückmeldungen ohne Automationsfehler statt und in der 31. Situation eine Rückmeldung mit Automationsfehler. Die Fahrsituationen waren anders verteilt als in der ersten Fahrthälfte (Abbildung 48). Die Situation mit Automationsfehler in der 62. Fahrsituation, unterschied sich nur optisch von der Situation mit Automationsfehler des 1. Fahrabschnitts. Das Vorderfahrzeug war in diesem Fall jedoch kein Auto, sondern ein Motorrad (Abbildung 45 rechts).



jedoch wurde diesmal in keiner Stufe des Faktors Rückmeldungsart eine Rückmeldung gegeben.

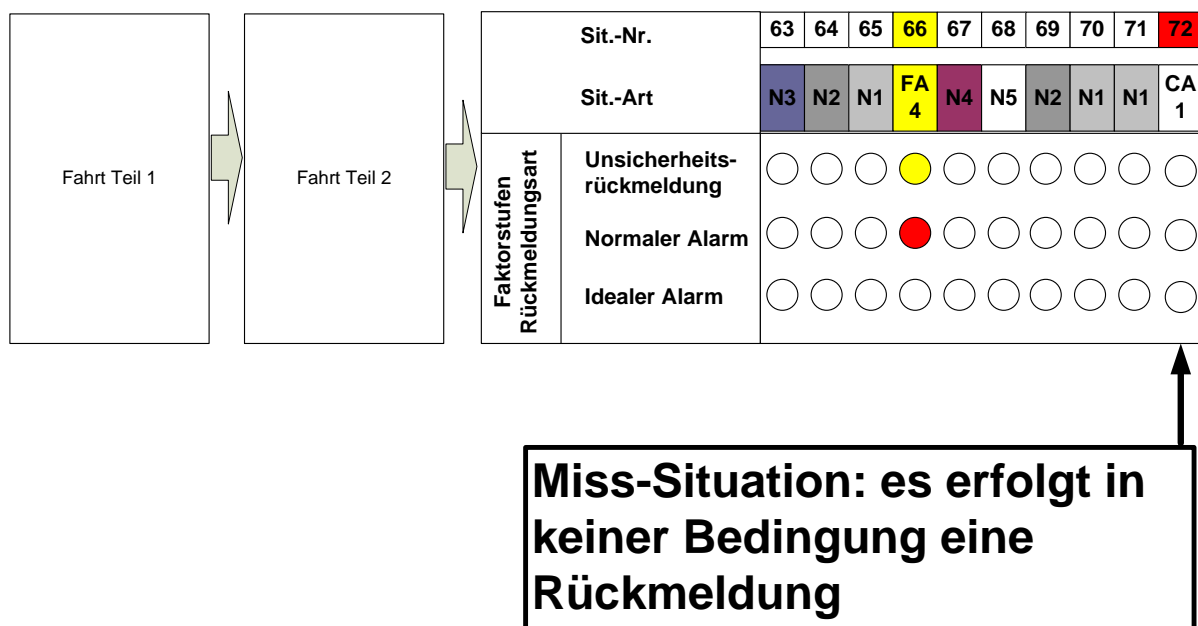
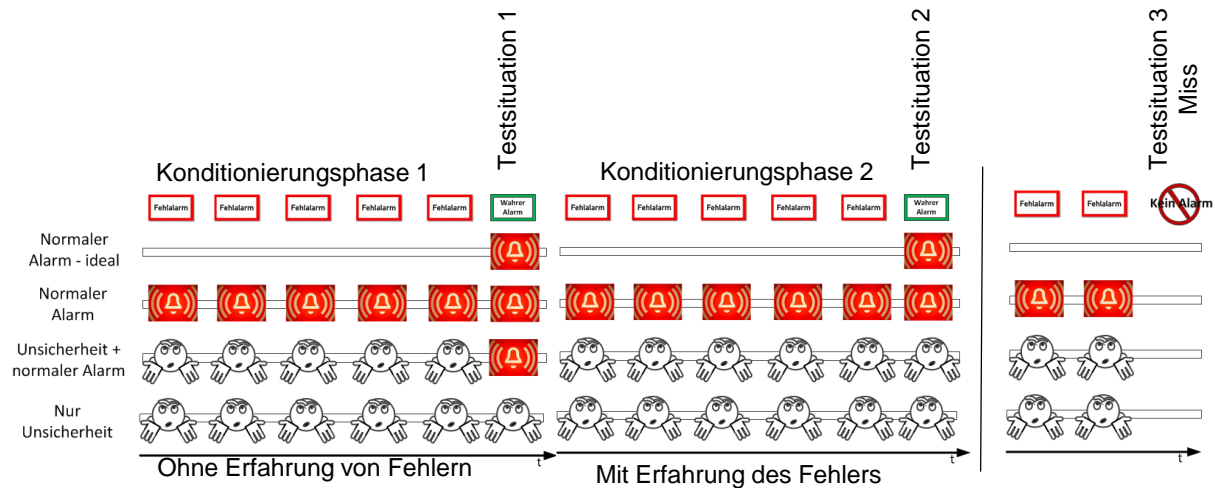


Abbildung 49: an die bereits bestehenden ersten Fahrteile angehangener dritter Fahrteil aus insgesamt 10 Fahrsituationen, der in der letzten Fahrsituation die „Miss-Situation“ enthielt.

### 9.2.3 Gesamtversuchsschema

Die zur Überprüfung der verschiedenen Hypothesen eingeführten Fahrten und Bedingungen wurden zu einem Gesamtversuchsschema kombiniert. Somit konnte jede Versuchsperson mehrere Teilerperimente in einer Sitzung bearbeiten. Zur besseren Verdeutlichung der resultierenden Fahrt- und Bedingungskombinationen dient Abbildung 50. Zusammenfassend wurde der Vergleich der Unsicherheitsrückmeldung mit anderen Alarmstrategien hinsichtlich der Cry Wolf Neigung in erster Linie im ersten Fahrteil durchgeführt. Die Überprüfung der Auswirkungen positiver Evidenz durch das Erleben eines auf die Rückmeldung folgenden Fehlers wurde durch einen Vergleich des ersten Fahrteils mit einem zweiten Fahrteil realisiert. Die Überprüfung der Auswirkungen einer neu eingeführten Bedingung Mischstrategie wurde ebenfalls mittels eines Vergleichs des ersten und zweiten Fahrteils durchgeführt. Die Überprüfung der Auswirkung unterschiedlicher Rückmeldungsarten auf die Kontrollierbarkeit in „Miss-Situationen“ wurde in einem zusätzlich hinzugefügten dritten Fahrteil durchgeführt.



**Abbildung 50: Schematischer Aufbau der 3. Studie.** Längs abgebildet sind die drei Versuchsfahrten, (von links nach rechts) erste Fahrt bis einschließlich der ersten Situation mit Automationsfehler (Testsituation 1), dann zweite Fahrt bis zur zweiten Situation mit Automationsfehler (Testsituation 2) und abschließend die dritte Fahrt mit Automationsfehler ohne Alarm. Untereinander abgebildet sind die unterschiedlichen Rückmeldungsarten / Strategien.

### 9.2.4 Vorgehen

Nachdem die Versuchspersonen eine Einverständniserklärung zur Datenaufzeichnung und einen demographischen Fragebogen ausgefüllt hatten, bekamen die Versuchspersonen eine Instruktion zum teil- und hochautomatisierten Fahren. Dies wurde getan, da die meisten Versuchspersonen keinerlei Vorstellung bzw. Erfahrung mit teil- oder hochautomatisiertem Fahren hatten. Nach der verbalen Instruktion bekamen die Versuchspersonen die Gelegenheit, sich in einem Training mit dem Fahren im Simulator und dem Fahren mit der Fahrzeugautomation vertraut zu machen. Für die Trainingsfahrt wurde das gleiche Grundszenario gewählt, welches auch in den späteren Versuchsfahrten verwendet wurde. Es wurden während der Fahrt jedoch nur klare Situationen ohne Rückmeldung oder Automationsfehler erlebt. In einem ersten 5minütigen Abschnitt fuhren die Versuchspersonen ohne Nebenaufgabe, in einem zweiten 5minütigen Abschnitt, sollten die Versuchspersonen die Bearbeitung der Nebenaufgabe während der Fahrt trainieren. Ebenso wurde die genaue Art der Nebenaufgabe und das Belohnungs-Bestrafungs- System erklärt (siehe Fahrerablenkung durch Nebenaufgaben). Das Training mit der Nebenaufgabe sollte mögliche Trainingseffekte im Verlauf der Versuchsfahrt minimieren. Nach dem Training wurde den Versuchspersonen die Automationsrückmeldung ihrer jeweiligen Rückmeldungsbedingung demonstriert und erklärt (siehe Abschnitt Umsetzung der Rückmeldungssituation).

Nach dieser Erklärung wurde die Versuchsfahrt begonnen. Die Versuchspersonen wurden instruiert, dass sie zu Beginn der Fahrt die Automation einschalten sollten, diese auf die Zielgeschwindigkeit von 100 km/h beschleunigen zu lassen und erst dann mit der Bearbeitung der Nebenaufgaben zu beginnen. Während der Fahrt erlebten die Versuchspersonen zunächst 30 unterschiedliche Fahrsituationen ohne Automationsfehler und, je nach Bedingung, in 5 dieser Situationen eine Automationsrückmeldung (Unsicherheitsrückmeldung in der Bedingung Unsicherheitsrückmeldung, normaler Alarm in der Bedingung normaler Alarm, keine Rückmeldung in der Bedingung idealer Alarm und Unsicherheitsrückmeldung in der Bedingung Mischstrategie). Nach jeder Fahrsituation sollten die Versuchspersonen während der Fahrt das situative Vertrauen in die Automation anhand einer 10 stufigen Skala einschätzen. In der 31. Fahrsituation wurde dann nach einer Rückmeldung ein Automationsfehler erlebt. In der Unsicherheitsbedingung wurde hier Unsicherheit zurückgemeldet, in der Bedingung normaler Alarm wieder ein normaler Alarm, in der Bedingung idealer Alarm wurde die erste Rückmeldung (normaler Alarm) erlebt und in der Bedingung Mischstrategie wurde ein normaler Alarm zurückgemeldet. Auch die Situation mit dem Automationsfehler wurde bezüglich des situativen Vertrauens bewertet.

Nach der Automationsfehlersituation wurden die Versuchspersonen gebeten, einfach weiter zu fahren und mit der Nebenaufgabenbearbeitung fortzufahren. Es folgten wieder 30 Fahrsituationen mit 5 unklaren Situationen. Diese waren in der zweiten Fahrt jedoch anders verteilt und unterschieden sich teilweise auch optisch. Ebenso im zweiten Fahrtteil wurde nach jeder Fahrsituation das situative Vertrauen erfragt. In der insgesamt 62. Fahrsituation geschah dann wieder ein Automationsfehler, vor dem eine Rückmeldung erfolgte. In der Bedingung Mischstrategie unterschied sich die Rückmeldung diesmal zu der beim ersten Automationsfehler, es wurde anstelle eines normalen Alarms nun Unsicherheit zurückgemeldet. In den anderen Bedingungen entsprach die Rückmeldung in der zweiten Automationsfehlersituation der Rückmeldung der ersten Fehlersituation.

Auch nach dieser Automationsfehlersituation sollten die Versuchspersonen die Fahrt fortsetzen. Es wurden noch 10 weitere Fahrsituationen durchfahren, von der die letzte Fahrsituation wieder eine Situation mit Automationsfehler war. In dieser Situation blieb die Automationsrückmeldung diesmal jedoch aus, um die „Miss-Situation“ abzubilden.

Nach dieser Situation wurde die Fahrt beendet. Die Versuchspersonen füllten noch einen abschließenden Fragebogen aus, wurden über den eigentlichen Zweck der Untersuchung aufgeklärt und erhielten die Aufwandsentschädigung für die Teilnahme am Versuch.



### 9.3 Ergebnisse Studie 3:

Da für die unterschiedlichen Teilerperimente häufig die gleichen abhängigen Variablen genutzt wurden, werden die Ergebnisse nicht je Hypothese / Teilerperiment, sondern je abhängiger Variable dargestellt.

#### 9.3.1 Stichprobe:

Die Versuchspersonen für Studie 3 wurden aus dem Probandenpool des DLR Instituts für Verkehrssystemtechnik akquiriert. Die 40 eingeladenen Versuchspersonen wurden, nach dem in Abschnitt „Generelles Vorgehen: Stichproben“ beschriebenen Parallelisierungsschema, auf die vier Versuchsgruppen aufgeteilt. Versuchspersonen wurden ausgeschlossen, wenn sie pauschal bei jeder Rückmeldung mit einer Bremsung reagierten (siehe Abschnitt: Generelles Vorgehen: Stichproben)

Um dennoch eine etwas größere Stichprobengröße als in Studie 2 zu erlangen, wurden so lange neue Versuchspersonen eingeladen, bis fast alle 4 Gruppen mit 10 Versuchspersonen besetzt werden konnten (Tabelle 32). Aufgrund der begrenzten Verfügbarkeit von Versuchspersonen konnte die Gruppe normaler Alarm nur mit 8 Versuchspersonen besetzt werden. Dennoch konnten die Parallelisierungskriterien weitestgehend eingehalten werden.

**Tabelle 32: Statistik der 4 Versuchspersonengruppen aus Studie 3**

Gruppe	N	Männlich / weiblich	Alter Min / Max	Alter Mittelwert / Stdabw	Führerschein Min / Max	Führerschein Mittelwert / Stdabw
Unsicherheitsrückmeldung	10	7 / 3	19 / 44	27,7 / 9,4	2 / 26	10,5 / 8,8
Normaler Alarm	8	5 / 3	19 / 51	33,3 / 12,8	3 / 33	15,8 / 12,7
Idealer Alarm	10	6 / 4	19 / 50	33,1 / 12,1	1 / 32	14,0 / 11,7
Mischstrategie	10	6 / 4	19 / 54	33,5 / 14,0	2 / 34	15,1 / 13,3

#### 9.3.2 Ergebnisse der Kontrollierbarkeit von Automationsfehlern

In diesem Abschnitt werden aus Gründen der Übersichtlichkeit für alle Hypothesen die Ergebnisse der Kontrollierbarkeit der Automationsfehler nacheinander in einzelnen Teilabschnitten berichtet.

Die Kontrollierbarkeit wird durch die Reaktionshäufigkeit, die Kollisionshäufigkeit und die durchschnittliche Bremsreaktionszeit operationalisiert. Auf die Berücksichtigung weiterer Maße zur Operationalisierung der Kontrollierbarkeit, wie etwa die minimum TTC oder die minimale Distanz zum Vorderfahrzeug, wird verzichtet, da sich die Bremsreaktionszeit als guter Indikator für die Güte der allgemeinen Reaktion erwies (siehe Abschnitt Messung des Cry Wolf Effektes).

Als Reaktionszeit wird in Studie 3 die Zeitdifferenz zwischen dem Erscheinen des Vorderfahrzeuges und dem ersten Betätigen des Bremspedals definiert. Sie beinhaltet dabei die Zeit, die Versuchspersonen benötigen zu erkennen, dass die Automation nicht auf das Vorderfahrzeug reagiert. Da die Geschwindigkeit des eigenen Fahrzeuges bis zum Betätigen des Bremspedals durch die Automation konstant gehalten wird, liegt zwischen dem Beginn einer Rückmeldung und dem Sichtbarwerden des Vorderfahrzeuges immer der gleiche Zeitraum.

### 9.3.2.1 Effekte der Rückmeldungsart auf die Kontrollierbarkeit des ersten erlebten Automationsfehlers in Situation 31 (ohne vorherige Fehlererfahrung – Hypothese zu Ziel 1 & Ziel 2)

Die Analyse der Reaktions- und Kollisionshäufigkeiten ergibt, dass in allen Rückmeldungsbedingungen, mit Ausnahme der Bedingung normaler Alarm, die Versuchspersonen mit einer Bremsung reagieren und auch erfolgreich eine Kollision vermeiden können.

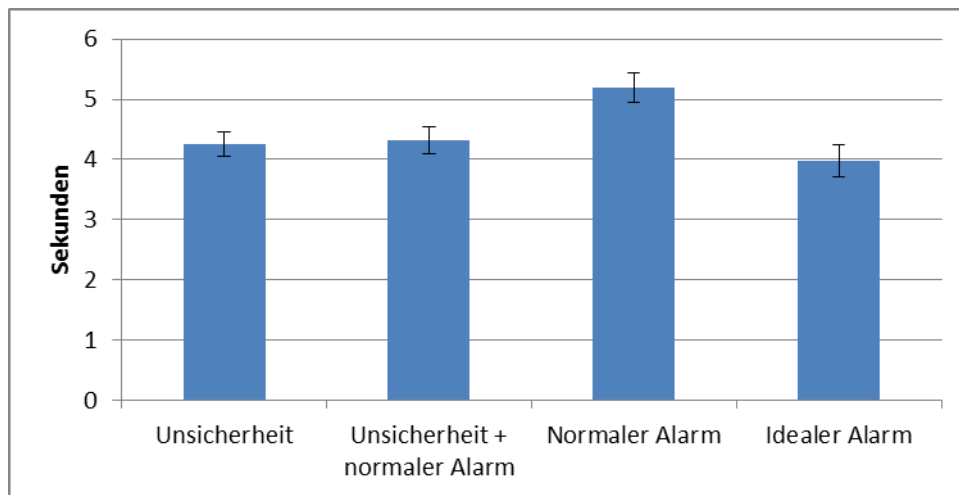
In der Bedingung normaler Alarm reagiert nur eine Versuchsperson nicht und kollidiert mit dem Vorderfahrzeug (Tabelle 33). Dies ist bereits ein augenscheinlicher Unterschied zu den beobachteten Reaktions- und Kollisionshäufigkeiten in Studie 2 und ein erster Hinweis darauf, dass in Studie 3 die Kontrolle der erwähnten Sekundärvarianz den gewünschten Effekt hat (siehe Methode Studie 3).

**Tabelle 33: Reaktions- und Kollisionshäufigkeiten in erster Automationsfehlersituation.**

	Bremsreaktionshäufigkeit (Anteil an Teilstichprobe)	Kollisionshäufigkeit (Anteil an Teilstichprobe)
<b>Unsicherheitsrückmeldung</b>	10 / 10	0 / 10
<b>Normaler Alarm</b>	7 / 8	1 / 8
<b>Idealer Alarm</b>	10 / 10	0 / 10
<b>Mischstrategie</b>	10 / 10	0 / 10

Bei der Betrachtung der Bremsreaktionszeiten zeigen sich zwischen den mittleren Reaktionszeiten der Rückmeldungsarten „Unsicherheit“, „idealer Alarm“ und der Kombination aus „Unsicherheit + normaler Alarm“ keine nennenswerten

Unterschiede (Tabelle 34 & Abbildung 51). Die Reaktionszeiten dieser drei Rückmeldungsarten unterscheiden sich aber signifikant bzw. tendenziell signifikant von der Reaktionszeit der Bedingung „normaler Alarm“ (siehe Tabelle 35). Die Reaktionszeit bei „normalem Alarm“ unterscheidet sich, wie ebenfalls angenommen, deutlich von der Reaktionszeit bei „idealem Alarm“ (1,21 Sekunden) und von der Kombination aus „Unsicherheit + normaler Alarm“ (0,87 Sekunden).



**Abbildung 51: Mittelwerte und Standardfehler für Reaktionszeit in Fehlersituation 1 (von dem Moment in dem das Vorderfahrzeug sichtbar wird bis zur ersten Betätigung des Bremspedals).**

**Tabelle 34: Deskriptive Statistik für Reaktionszeit in erster Fehlersituation**

Alarm	Mittelwert	Standardabweichung	N
Normaler Alarm	5,18988	,635994	8
Idealer Alarm	3,97960	,827351	10
Unsicherheit	4,25450	,660794	10
Unsicherheit + normaler Alarm	4,31910	,700947	10

Die ANOVA ergibt einen signifikanten Haupteffekt des Faktors Alarm  $F(3, 34) = 4,597$  mit einer mittleren Effektstärke ( $\eta^2 = .289$  partielles Eta Quadrat).

Auf Ebene der paarweisen Vergleiche (Tabelle 35) zeigt sich, wie erwartet, ein signifikanter Unterschied der Reaktionszeit zwischen der Unsicherheitsrückmeldung und dem normalen Alarm. Die Reaktionszeit bei Unsicherheitsrückmeldung liegt im Schnitt eine Sekunde unter der Reaktionszeit bei normalem Alarm (0,935 Sekunden). Die Effektstärke beim paarweisen Vergleich ist mittelgroß, mit einer Tendenz zu einem starken Effekt (partielles Eta  $\eta^2 = .0365$ ).

Hingegen gibt es keine, auch nicht tendenziell signifikanten Unterschiede zwischen der Unsicherheitsrückmeldung und dem idealen Alarm (0,27 Sekunden), bzw. der

Kombination aus Unsicherheitsrückmeldung und normalem Alarm (0,06 Sekunden). Diese Beobachtungen entsprechen den formulierten Hypothesen.

**Tabelle 35: p Werte der paarweisen Vergleiche (Korrekturverfahren - Tukey HSD)**

	<b>Normaler Alarm</b>	<b>Idealer Alarm</b>	<b>Unsicherheit</b>
<b>Idealer Alarm</b>	.006		
<b>Unsicherheit</b>	.044	.825	
<b>Unsicherheit + normaler Alarm</b>	.067	.714	.997

### **Effekt positiver Evidenz in erster Fehlersituation auf die Entstehung eines Cry Wolf Effektes zur zweiten Fehlersituation.**

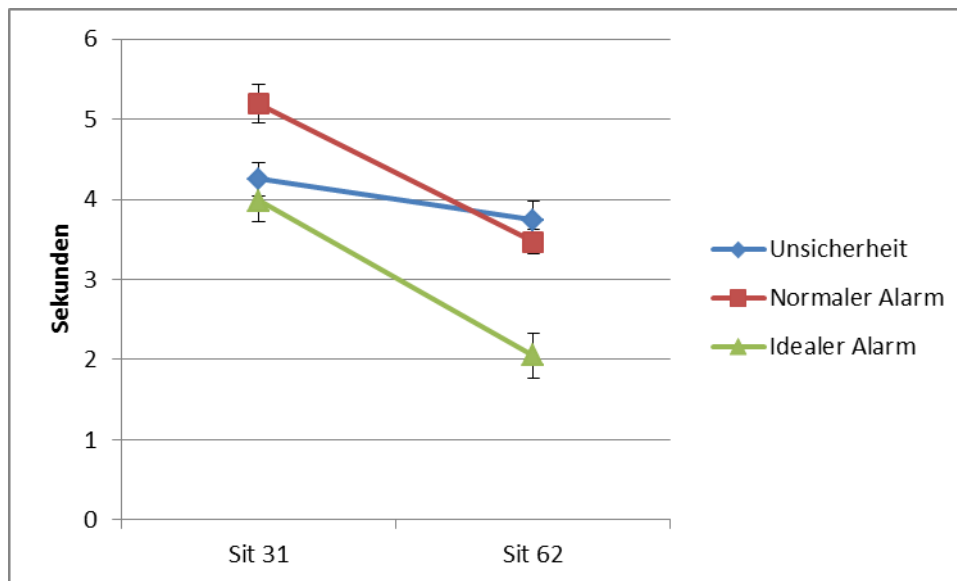
Betrachtet werden die Rückmeldungsarten Unsicherheit, normaler Alarm und idealer Alarm (Hypothese zu Ziel 3).

Im Folgenden werden die durchschnittlichen Reaktionszeiten zwischen der ersten und zweiten Automationsfehlersituation verglichen. Die Reaktionszeitergebnisse der Bedingung Mischstrategie werden in einem separaten Abschnitt vorgestellt.

Im Gegensatz zur ersten Automationsfehlersituation können in der zweiten Automationsfehlersituation alle Versuchspersonen eine Kollision vermeiden. Im Vergleich zu den Reaktionszeiten der ersten Automationsfehlersituation liegen die beobachteten Reaktionszeiten insgesamt unter denen der ersten Fehlersituation. Die Reaktionszeit in der Bedingung idealer Alarm liegt in der zweiten Fehlersituation nun deutlich unter den Reaktionszeiten der Bedingungen Unsicherheit und normaler Alarm (Abbildung 52). Ebenso liegt die Reaktionszeit der Bedingung normaler Alarm diesmal unterhalb der Reaktionszeit bei Unsicherheitsrückmeldung (Tabelle 36).

**Tabelle 36: Deskriptive Statistik der Reaktionszeit in der zweiten Automationsfehlersituation**

<b>Alarm</b>	<b>Mittelwert</b>	<b>Standardabweichung</b>	<b>N</b>
<b>Normaler Alarm</b>	3,47	,405	7
<b>Idealer Alarm</b>	2,04	,874	10
<b>Unsicherheit</b>	3,74	,726	10



**Abbildung 52: Interaktion zwischen Rückmeldungsart und Erfahrung von Fehlerereignissen.**

Eine Mixed Models Varianzanalyse des Innergruppenfaktors Evidenzerfahrung und des Zwischensubjektfaktors Rückmeldungsart zeigt einen deutlich signifikanten ( $F_{46,347} df=1, 24 p = .000$ ) Effekt der Fehlererfahrung mit einer hohen Effektstärke (partielles Eta Quadrat .659).

Ebenso kann ein Interaktionseffekt ( $F_{5,249} df = 2,24 p = .013$ ) zwischen der Rückmeldungsart und der Fehlererfahrung mit mittlerer Effektstärke (partielles Eta Quadrat .304) festgestellt werden. In Abhängigkeit von der Rückmeldungsart, wirkt sich die Fehlererfahrung also unterschiedlich aus.

Erwähnenswert ist auch, dass in der Bedingung normaler Alarm, im zweiten Fahrtabschnitt, 2 Versuchspersonen in zusammen 5 Situationen mit korrektem Automationsverhalten mit einer Bremsung auf den Alarm reagieren. Im ersten Fahrtabschnitt konnte dies nur in einer Situation bei einer Versuchsperson beobachtet werden. Bezeichnenderweise hat eine dieser Versuchspersonen eine der schlechtesten Reaktionszeiten im ersten Fahrtabschnitt (7,74 sek) und die andere Versuchsperson reagierte als einzige in der Fehlersituation des ersten Fahrtabschnitts gar nicht auf den Alarm und kollidierte. Für beide Versuchspersonen ist demnach von einem ausgeprägten Cry Wolf Effekt im ersten Fahrtabschnitt auszugehen, der dazu führt, dass im zweiten Fahrtabschnitt deutlich sensibler auf den Alarm reagiert wird. In der Bedingung Unsicherheitsrückmeldung zeigt sich hingegen, dass in der ersten Fahrt zwei Versuchspersonen in je einer Situation mit einer leichten Bremsung reagieren, in der zweiten Fahrt jedoch nur eine Versuchsperson in einer Situation.

Diese Ergebnisse stützen die Hypothese, dass die Erfahrung positiver Evidenz, sich in erster Linie bei einem normalen Alarm auswirkt.

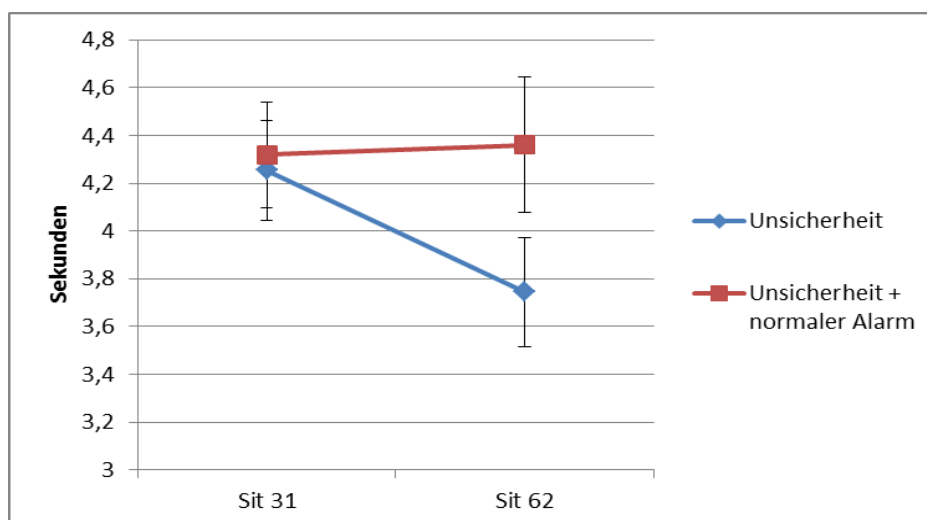
### 9.3.2.2 Effekt positiver Evidenz in erster Fehlersituation auf die Entstehung eines Cry Wolf Effektes zur zweiten Fehlersituation für die Rückmeldungsart Mischstrategie (Hypothese zu Ziel 4)

Für diese Betrachtung werden die Reaktionszeiten der Bedingung „Mischstrategie“ in der ersten und zweiten Fehlersituation, jeweils mit den Reaktionszeiten der Bedingung „Unsicherheit“ verglichen (Abbildung 53 & Tabelle 37). In der Bedingung Unsicherheit gibt es einen augenscheinlichen Unterschied zwischen der Reaktionszeit in der 1. und 2. Automationsfehlersituation. In der Bedingung Mischstrategie findet jedoch kaum eine Veränderung statt. Die Überprüfung mittels Mixed Models Varianzanalyse ergibt jedoch keinen signifikanten Interaktionseffekt zwischen Rückmeldungsart (Unsicherheit & Kombination) und Fehlererfahrung. ( $F_{2,098} Df = (1,17) p. 166$ ).

**Tabelle 37: Bremsreaktionszeiten der Bedingungen Unsicherheitsrückmeldung und Mischstrategie in 1. & 2. Automationsfehlersituation**

	Automationsfehler 1	Automationsfehler 2
<b>Unsicherheit</b>	Unsicherheit 4,254 sek.	Unsicherheit 3,743 sek.
<b>Mischstrategie</b>	Normaler Alarm 4,307 sek.	Unsicherheit 4,360 sek.

Die Ergebnisse lassen sich dennoch in der Richtung deuten, dass bei einer Mischstrategie das Verwenden einer Rückmeldung Variante A als Regelfall bei einem Automationsfehler, die Effektivität einer Rückmeldung Variante B als Ausnahme bei einem Automationsfehler, negativ beeinflussen kann.



**Abbildung 53: Reaktionszeiten der Bedingungen „Unsicherheit“ und „Mischstrategie“ (normaler Alarm + Unsicherheit) in erster und zweiter Fehlersituation.**

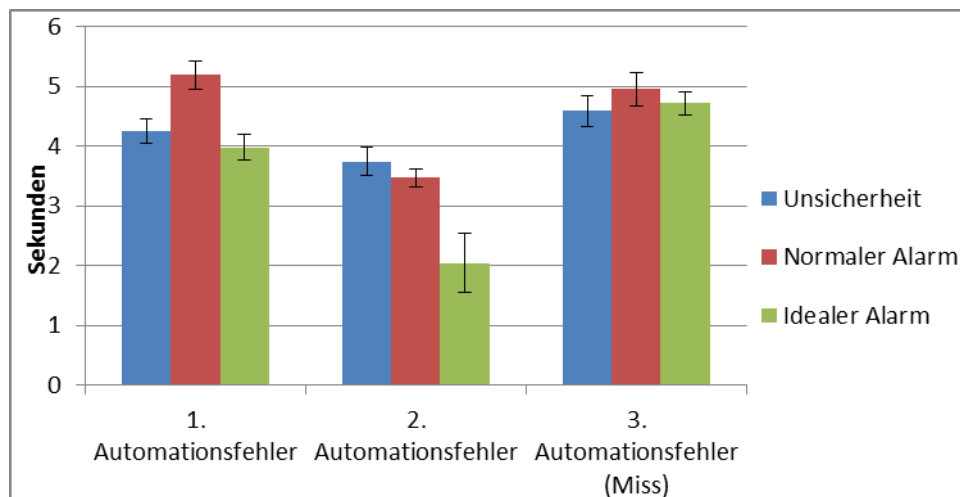
### 9.3.3 Ergebnisse der Reaktionszeit in der Miss-Situation

Verglichen werden die Bremsreaktionszeiten in den Rückmeldungsbedingungen Unsicherheit, normaler Alarm und idealer Alarm.

In der Miss-Situation zeigen zwar alle Versuchspersonen eine Bremsreaktion, jedoch kommt es, im Gegensatz zu den vorherigen Automationsfehlersituationen, bei einigen Versuchspersonen zu Kollisionen mit dem Vorderfahrzeug (Tabelle 38). Bei der Betrachtung der Reaktionszeiten liegen die durchschnittlichen Bremsreaktionszeiten der Rückmeldungsarten auf einem vergleichbaren Niveau (Abbildung 54). Eine durchgeführte ANOVA kann entsprechend keinen Unterschied zwischen den Rückmeldungsbedingungen zeigen ( $p = .76$ ).

**Tabelle 38: Bremsreaktions- und Kollisionshäufigkeit in der Miss-Situation**

	Bremsreaktionshäufigkeit (Anteil an Teilstichprobe)	Kollisionshäufigkeit (Anteil an Teilstichprobe)
Unsicherheitsrückmeldung	10 / 10	2 / 10
Normaler Alarm	7 / 7	1 / 7
Idealer Alarm	10 / 10	1 / 10



**Abbildung 54: Reaktionszeit in Miss Situation im Vergleich zu Reaktionszeiten bei Automationsfehlern in Situationen mit Rückmeldung**

Die Bremsreaktionszeiten in der Miss-Situation liegen, mit Ausnahme der Bedingung normaler Alarm in der ersten Automationsfehlersituation, insgesamt deutlich höher als in den vorherigen Automationsfehlersituationen. Die für den Vergleich der verschiedenen Automationsfehlersituationen durchgeführte Messwiederholungs-

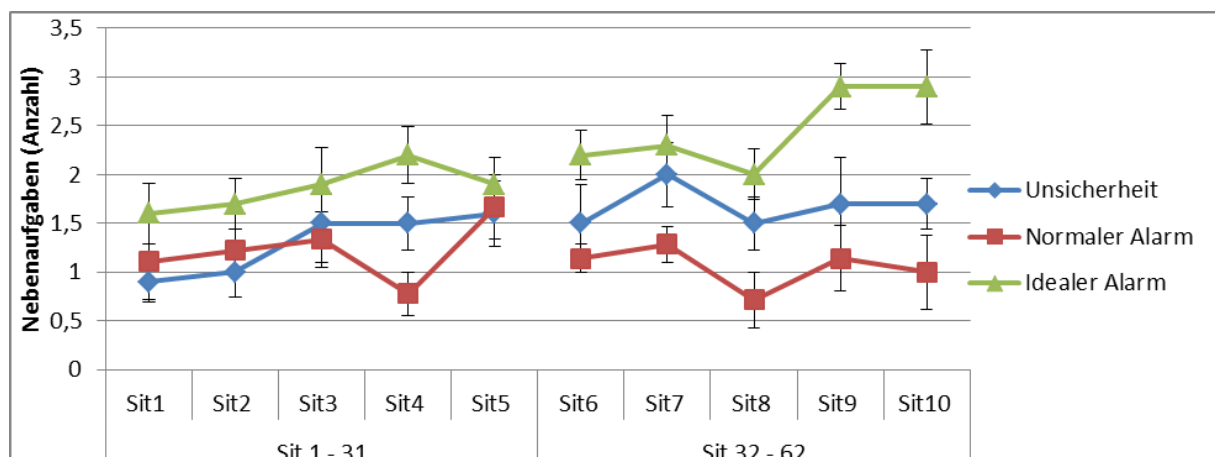
Varianzanalyse ergibt entsprechend einen signifikanten Haupteffekt ( $F_{12,012}$ ,  $df = 2/48$ ,  $p .000$ , partielles Eta  $.334$ ).

Insgesamt scheint die mit den jeweiligen Rückmeldungsarten gemachte Vorerfahrung in den Situationen vor der Miss-Situation, keinen nachweisbaren Einfluss auf die Reaktionsqualität in der Miss-Situation zu haben.

In der Bedingung normaler Alarm ist die vergleichbar niedrige Reaktionszeit in der ersten Automationsfehlersituation und in der Miss-Situation ein Hinweis darauf, dass in der ersten Automationsfehlersituation der normale Alarm keinerlei Effekt mehr gehabt zu haben scheint. Es wurde in der Fehlersituation mit Rückmeldung eine ähnlich schlechte Performanz erreicht, wie in einer Situation ohne Rückmeldung.

### 9.3.4 Ergebnisse Nebenaufgabenbearbeitung

Zur Analyse der Entwicklung eines Cry Wolf Effektes in Abhängigkeit der Rückmeldungsart und der Erfahrung eines Automationsfehlers, wird der Grad der Hinwendung zu einer Nebenaufgabe in den Situationen einer Automationsrückmeldung ohne Automationsfehler betrachtet. Dazu wird, je Fahrtabschnitt, die Entwicklung der Nebenaufgabenleistung über die fünf Rückmeldungssituationen in Abhängigkeit der Rückmeldungsarten Unsicherheit und normaler Alarm betrachtet. Als Referenzbedingung dient die Bedingung idealer Alarm, da in dieser Bedingung die Nebenaufgabenleistung nicht durch eine Rückmeldung beeinflusst wird. Ebenso wird die durchschnittliche Nebenaufgabenleistung für jede Rückmeldungsart zwischen den Fahrtabschnitten verglichen.



**Abbildung 55: Entwicklung der Nebenaufgabenleistung in den Rückmeldesituationen ohne Automationsfehler über die gesamte Fahrt**



#### 9.3.4.1 Entwicklung der Nebenaufgabenleistung innerhalb Fahrtabschnitt 1

Im ersten Fahrtabschnitt liegt in der Bedingung idealer Alarm die Nebenaufgabenleistung in allen Fahrsituation über der Nebenaufgabenleistung der Bedingungen Unsicherheit und normaler Alarm. Es zeigt sich jedoch kein Haupteffekt des Faktors Rückmeldungsart, auch paarweise Vergleiche der Rückmeldungsarten können keinen bedeutsamen Unterschied zwischen den Rückmeldungsarten zeigen (Tabelle 40).

Über die Fahrt steigt die Nebenaufgabenleistung in allen Rückmeldungsbedingungen leicht, aber signifikant an. Es gibt einen Haupteffekt des Innergruppenfaktors Situationsabfolge (Tabelle 39).

Es kann kein Interaktionseffekt aus Situationsabfolge und Rückmeldungsart gefunden werden. Dies weist darauf hin, dass sich der Anstieg zwischen den Rückmeldungsgruppen in einem ähnlichen Maße vollzieht. Erwartet wurde jedoch, dass sich die Nebenaufgabenleistung in der Bedingung normaler Alarm etwas stärker erhöht.

**Tabelle 39: Ergebnisse der Messwiederholungs-Varianzanalyse für den ersten Fahrtabschnitt (df korrigiert nach Greenhouse-Geisser):**

Faktor	F	df	Sig.	Eta
Rückmeldungsart	2,101	2 / 26	,143	,139
Situationsabfolge	3,013	6,490 / 104	,031	,104
Rückmeldungsart x Reihenfolge des Automationsfehlers	1,515	3,245 / 104	,178	,104

**Tabelle 40: Paarweise Vergleiche der Rückmeldungsbedingungen im ersten Fahrtabschnitt - p Werte (korrigiert mit TukeyHSD)**

	Unsicherheit	Normaler Alarm
Normaler Alarm	,972	
Idealer Alarm	,236	,173

#### 9.3.4.2 Entwicklung der Nebenaufgabenleistung innerhalb Fahrtabschnitt 2

Im zweiten Fahrtabschnitt gibt es insgesamt einen deutlichen Haupteffekt des Faktors Rückmeldungsart (Tabelle 41). Deskriptiv unterscheidet sich die Nebenaufgabenleistung in den einzelnen Rückmeldungsbedingungen über den gesamten Fahrtverlauf hinweg. Die Nebenaufgabenleistung in der Bedingung normaler Alarm scheint nach dem Erleben einer positiven Evidenz, wie erwartet,

durchweg geringer zu sein als die der Bedingung Unsicherheit. Hier konnte im ersten Fahrtabschnitt kein systematischer Unterschied beobachtet werden.

Die paarweisen Vergleiche ergeben jedoch nur einen signifikanten Unterschied zwischen der Bedingung idealer Alarm und normaler Alarm. Der Unterschied zwischen idealem Alarm und Unsicherheit ist nur tendenziell signifikant und zwischen Unsicherheit und normalem Alarm kann kein bedeutsamer Unterschied festgestellt werden (Tabelle 42).

Insgesamt gibt es einen signifikanten Effekt des Innergruppenfaktors Situationsabfolge (Tabelle 41), wobei jedoch ein deutlicher Anstieg über die Fahrt nur in der Bedingung idealer Alarm beobachtet werden kann. Es wird daher eine separate Analyse ohne die Bedingung idealer Alarm durchgeführt, die jedoch keinen Effekt der Situationsabfolge bestätigt (Tabelle 43).

Es kann auch keine Interaktion zwischen der Situationsabfolge und den Rückmeldungsbedingungen Unsicherheit und normaler Alarm gezeigt werden. D.h., es ist davon auszugehen, dass sich die Nebenaufgabenleistung in beiden Rückmeldungsarten über die Fahrt nicht systematisch erhöht.

**Tabelle 41: Ergebnisse der Messwiederholungs-Varianzanalyse für den zweiten Fahrtabschnitt (df korrigiert nach Greenhouse-Geisser):**

Faktor	F	df	Sig.	Eta
Rückmeldungsart	6,673	1/ 24	,005	,357
Situationsabfolge	3,303	3,06 / 73,58	,024	,121
Rückmeldungsart x Reihenfolge des Automationsfehlers	1,489	6,132 / 73,58	,193	,110

**Tabelle 42: Paarweise Vergleiche der Rückmeldungsbedingungen im zweiten Fahrtabschnitt - p Werte (korrigiert mit TukeyHSD)**

	Unsicherheit	Normaler Alarm
Normaler Alarm	,266	
Idealer Alarm	,091	,004

Tabelle 43: Ergebnisse der Messwiederholungs-Varianzanalyse für den zweiten Fahrtabschnitt, nur für die Bedingungen Unsicherheit und normaler Alarm (df korrigiert nach Greenhouse-Geisser).

Faktor	F	df	Sig.	Eta
Rückmeldungsart	2,288	1/ 15	,151	,132
Situationsabfolge	1,719	2,064 / 30,962	,195	,103
Rückmeldungsart x Reihenfolge des Automationsfehlers	,334	2,064 / 30,962	,725	,022

#### 9.3.4.3 Vergleich der Nebenaufgabenleistung zwischen Fahrtabschnitt 1 und 2

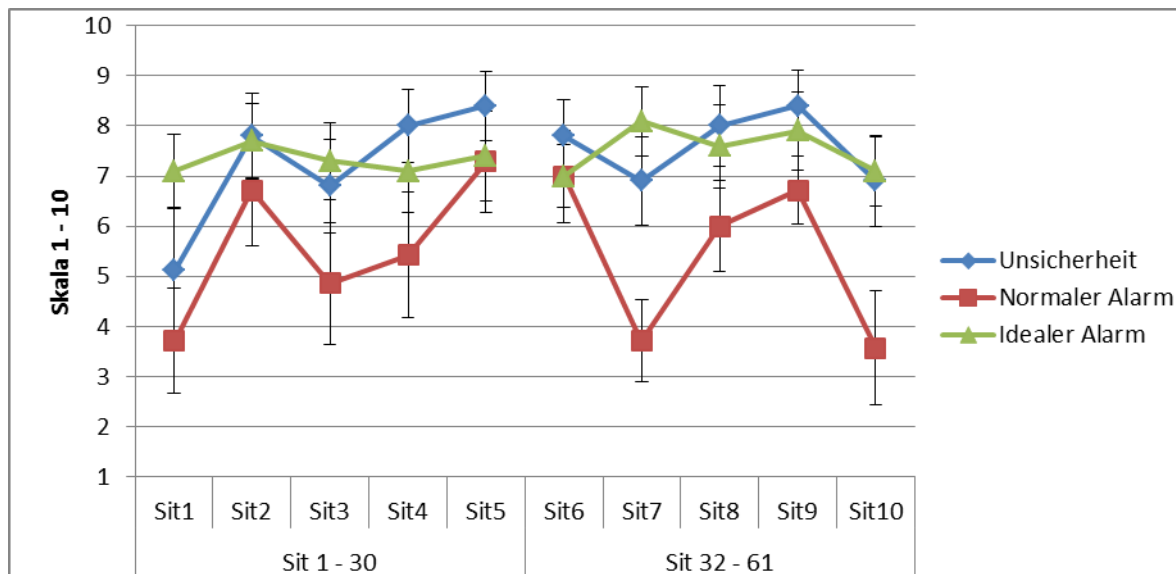
Die Reihenfolge der Automationsfehlersituation hat einen signifikanten Haupteffekt, d.h. die Nebenaufgabenleistung im zweiten Automationsfehler unterscheidet sich deutlich von der in der ersten Automationsfehlersituation (Tabelle 44). Es zeigt sich ein signifikanter Interaktionseffekt der Faktoren Rückmeldungsart und der Situationsabfolge. Die Erfahrung eines Automationsfehlers in der ersten Situation mit Automationsfehler wirkt sich also in Abhängigkeit der Rückmeldungsart unterschiedlich auf die Nebenaufgabenleistung im zweiten Fahrtabschnitt aus.

Tabelle 44: Ergebnisse der Messwiederholungs-Varianzanalyse mit Innergruppenfaktor Reihenfolge der Fahrtabschnitte (-Automationsfehlers) und Zwischengruppenfaktor Rückmeldungsart für die drei Stufen Unsicherheit, normaler Alarm und Idealer Alarm.

Faktor	F	df	Sig.	Eta
Reihenfolge des Automationsfehlers	7,500	1 / 24	<b>,011</b>	,238
Rückmeldungsart x Reihenfolge des Automationsfehlers	6,647	2 / 24	<b>,005</b>	,356

#### 9.3.5 Ergebnisse situatives Vertrauen

Zur weiteren Analyse der Auswirkungen einer Unsicherheitsrückmeldung wird das situative Vertrauen in die Automation unmittelbar nach Rückmeldungssituationen zwischen normalem Alarm und Unsicherheitsrückmeldung verglichen (siehe Erfassung des situativen Vertrauens). Als Referenzbedingung dient wieder die Bedingung idealer Alarm, da in dieser Bedingung das situative Vertrauen nicht durch eine Rückmeldung beeinflusst wird.



**Abbildung 56: Entwicklung des situativen Vertrauens nach Rückmeldungssituationen ohne Automationsfehler, je im ersten Fahrtabschnitt (links) und im zweiten Fahrtabschnitt (rechts).**

### 9.3.5.1 Entwicklung des situativen Vertrauens innerhalb Fahrtabschnitt 1

In der Bedingung idealer Alarm verändert sich das situative Vertrauen augenscheinlich nicht über die 5 Rückmeldungssituationen (Abbildung 56). In den Bedingungen Unsicherheit und normaler Alarm scheint es einen systematischen Anstieg über die 5 Rückmeldungssituationen zu geben. Der Faktor Situationsabfolge hat einen signifikanten Haupteffekt (Tabelle 45). Jedoch gibt es weder einen Interaktionseffekt aus Rückmeldungsart und Situationsabfolge noch einen Haupteffekt des Faktors Rückmeldung. Auch die paarweisen Vergleiche (Tabelle 46) zeigen keine Unterschiede zwischen den Rückmeldungsarten.

**Tabelle 45: Ergebnisse der Messwiederholungs-Varianzanalyse für den ersten Fahrtabschnitt**

Faktor	F	df	Sig.	Eta
Rückmeldungsart	1,508	1 / 23	,242	,116
Situationsabfolge	5,568	2,490 / 57,278	<b>,004</b>	,195
Rückmeldungsart x Situationsabfolge	1,251	4,981 / 57,278	,298	,098

**Tabelle 46: Paarweise Vergleiche der Rückmeldungsbedingungen im ersten Fahrtabschnitt - p Werte (korrigiert mit TukeyHSD)**

	Unsicherheit	Normaler Alarm
Normaler Alarm	,353	
Idealer Alarm	,975	,249

### 9.3.5.2 Entwicklung des situativen Vertrauens innerhalb Fahrtabschnitt 2

Im zweiten Fahrtabschnitt steigt in der Bedingung Unsicherheit das situative Vertrauen nicht weiter über die Fahrt an, sondern bleibt auf einem konstant hohen Niveau, in etwa vergleichbar mit dem situativen Vertrauen der Bedingung idealer Alarm (Abbildung 56). Innerhalb der Bedingung normaler Alarm findet eine intensive Schwankung des situativen Vertrauens über die 5 Rückmeldesituationen des zweiten Fahrtabschnittes statt. Deskriptiv liegt das situative Vertrauen insgesamt unter dem der beiden anderen Rückmeldungsbedingungen. Es gibt aber nur einen tendenziell signifikanten Haupteffekt des Faktors Rückmeldung (Tabelle 47). Auch bei den paarweisen Vergleichen der Rückmeldungsbedingungen ist sowohl der Unterschied zwischen idealem Alarm und normalen Alarm als auch der Unterschied zwischen Unsicherheit und normalem Alarm nicht signifikant (Tabelle 48).

**Tabelle 47: Ergebnisse der Messwiederholungs-Varianzanalyse für den zweiten Fahrtabschnitt**

Faktor	F	df	Sig.	Eta
Rückmeldungsart	2,866	1 / 24	,077	,193
Situationsabfolge	6,867	2,279 / 54,7	,001	,222
Rückmeldungsart x Situationsabfolge	3,033	4,558 / 54,7	,020	,202

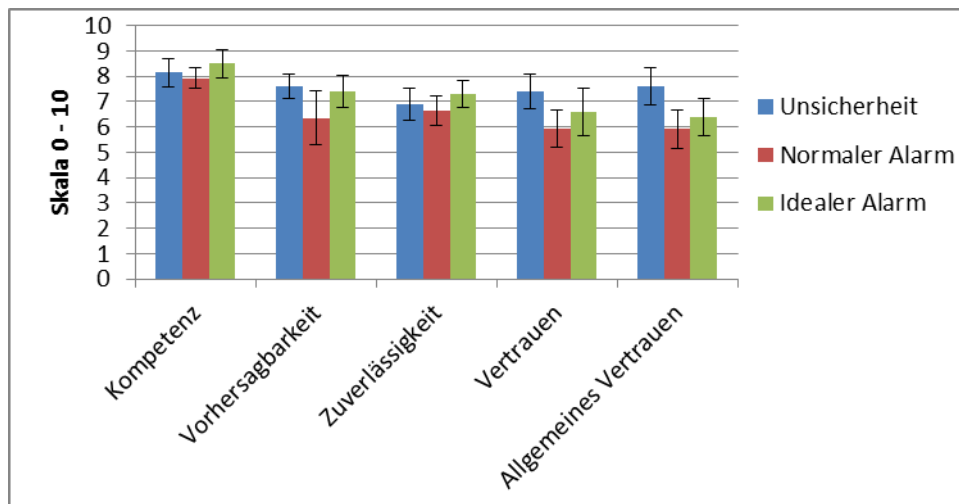
**Tabelle 48: Paarweise Vergleiche der Rückmeldungsbedingungen im zweiten Fahrtabschnitt - p Werte (korrigiert mit TukeyHSD)**

	Unsicherheit	Normaler Alarm
Normaler Alarm	,099	
Idealer Alarm	,998	,110

### 9.3.6 Ergebnisse Gesamtvertrauen in die Automation

#### 9.3.6.1 Ergebnisse des Gesamtvertrauens nach dem 2. Automationsfehler.

Das Gesamtvertrauen wurde durch den bereits in Studie 1 & 2 verwendeten Trustfragebogen zwischen der zweiten Automationsfehlersituation und vor dem Fahrtabschnitt mit den Miss-Situationen erhoben.



**Abbildung 57: Vertrauen in die Automation insgesamt nach der zweiten Automationsfehlersituation (Sit. 62)**

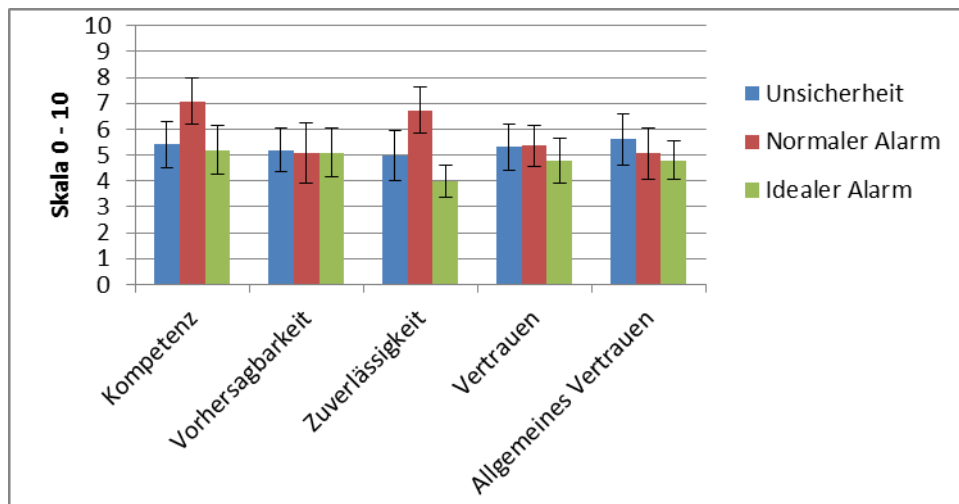
Eine multivariate Varianzanalyse des Faktors Rückmeldung ergibt weder einen Effekt bei multivariatem Test unter Einbeziehung aller 5 Trustdimensionen (Pillai-Spur Wert = .240  $F = .530$   $df\ 10 / 42$   $p = .827$  partielles Eta = .120), noch einen signifikanten Unterschied zwischen den Rückmeldungsarten auf den einzelnen Trustdimensionen (Tabelle 49).

**Tabelle 49: Ergebnisse der ANOVA des Faktors Rückmeldung für die 5 Dimensionen des Trustfragebogens.**

	Trustdimension	df	F	Sig.	Partielles Eta
Rückmeldung	Kompetenz	2,24	.329	.723	,027
	Vorhersagbarkeit	2,24	.693	.510	,055
	Zuverlässigkeit	2,24	.240	.788	,020
	Vertrauen	2,24	.676	.518	,053
	Vertrauen allgemein	2,24	1.210	.316	,092

### 9.3.6.2 Gesamttrust insgesamt über alle Situationen inklusive Miss-Situation

Nach der letzten Fahrsituation, der Miss-Situation, wurde nochmals das Vertrauen in die Automation erhoben.



**Abbildung 58: Ergebnisse des Fragebogens Vertrauen in die Automation bei Erfassung unmittelbar nach dem Erleben der „Miss-Situation“**

Eine multivariate Varianzanalyse der 5 Trustdimensionen ergibt keinen Haupteffekt des Faktors Rückmeldungsart (Pillai-Spur Wert = .352  $F = .897$   $df\ 10 / 42$   $p = .544$  partielles Eta = .176). Auch gibt es keine signifikanten Unterschiede zwischen den Rückmeldungsarten auf den jeweilig einzelnen Trustdimensionen (Tabelle 50).

**Tabelle 50: Ergebnisse der ANOVA des Faktors Rückmeldung für die 5 Dimensionen des Trustfragebogens.**

	Trustdimension	df	F	Sig.	Partielles Eta
Rückmeldung	Kompetenz	2,24	.980	.390	.075
	Vorhersagbarkeit	2,24	.010	.990	.001
	Zuverlässigkeit	2,24	2.411	.111	.167
	Vertrauen	2,24	.112	.895	.009
	Vertrauen allgemein	2,24	.225	.800	.018

### 9.3.6.3 Vergleich des Vertrauens in die Automation vor und nach der Miss-Situation

Die Beurteilungen des Gesamtvertrauens nach dem zweiten Fahrtabschnitt werden mit denen nach der Miss-Situation (3. Fahrtabschnitt) verglichen. Die multivariate Varianzanalyse über alle 5 Trust-Dimensionen zeigt einen Haupteffekt der Reihenfolge. (Pillai-Spur Wert = .503  $F = .4.044$   $df\ 5 / 20$   $p = .011$  partielles Eta = .503). Es gibt keinen Effekt der Rückmeldungsart und auch keinen Interaktionseffekt. Bei Betrachtung der einzelnen Trust-Dimensionen zeigt sich das gleiche Bild. Auch auf allen einzelnen Trustdimensionen gibt es einen signifikanten Haupteffekt der Situationsabfolge (Tabelle 51) und keinen Effekt der Rückmeldungsart. Die

Rückmeldungsart scheint demzufolge keinen moderierenden Einfluss auf die Wirkung der Miss-Situation auf das Vertrauen in die Automation zu haben. Das Erleben der Miss-Situation an sich hat deutlich negative Auswirkungen auf das Vertrauen in die Automation.

**Tabelle 51: Ergebnisse der Messwiederholungsanalyse des Faktors Situationsabfolge für die 5 Dimensionen des Trustfragebogens.**

	Trustdimension	df	F	Sig.	Partielles Eta
Situationsabfolge	Kompetenz	1,24	17.410	.000	.420
	Vorhersagbarkeit	1,24	17.982	.000	.428
	Zuverlässigkeit	1,24	14.232	.001	.372
	Vertrauen	1,24	11.650	.002	.327
	Vertrauen allgemein	1,24	9.593	.005	.286

**Tabelle 52: Ergebnisse der Messwiederholungsanalyse für den Zwischengruppenfaktor Rückmeldung für die 5 Dimensionen des Trustfragebogens.**

	Trustdimension	df	F	Sig.	Partielles Eta
Rückmeldung	Kompetenz	2,24	.319	.730	.026
	Vorhersagbarkeit	2,24	.212	.811	.017
	Zuverlässigkeit	2,24	.703	.505	.055
	Vertrauen	2,24	.276	.762	.022
	Vertrauen allgemein	2,24	.724	.495	.057

## 9.4 Diskussion Studie 3

Studie 3 sollte die Bearbeitung unterschiedlicher Forschungsfragen ermöglichen und wurde daher so aufgebaut, unterschiedliche Ziele realisieren zu können. Die Diskussion der Ergebnisse wird in der Reihenfolge der in Abschnitt Ziele Studie 3 aufgeführten Ziele gegliedert.

**Ziel 1:** Wiederholung des Vergleichs von Unsicherheitsrückmeldung und normalem Alarm:

In Studie 2 wurde bereits ein Vergleich zwischen Unsicherheitsrückmeldung und einem konventionell gestalteten Alarm durchgeführt. Es konnten dabei keine eindeutig signifikanten Unterschiede identifiziert werden. Es wird vermutet, dass durch die gleichartige Gestaltung der Situationen ohne Automationsfehler und der Situation mit Automationsfehler ein Lerneffekt stattfand, der den Cry Wolf Effekt



konfundierte (zur genaueren Erläuterung siehe Diskussionsteil Studie 2). Ausgehend von dieser Vermutung sollte dieser konfundierende Lerneffekt in Studie 3 durch stärkere Variation des Erscheinungsbildes der Rückmeldungssituationen minimiert werden. Die Ergebnisse der Kontrollierbarkeitssituation des ersten Fahrtabschnitts (Situation 1 bis 31) zeigen in Studie 3 nun einen deutlicheren Unterschied zwischen der Unsicherheitsrückmeldung und dem normalen Alarm als noch in Studie 2. In der Unsicherheitsbedingung können, im Gegensatz zu Studie 2, alle Versuchspersonen erfolgreich eine Kollision vermeiden und alle Versuchspersonen zeigen eine Bremsreaktion. Die Reaktionszeit fällt im Durchschnitt signifikant niedriger aus, als in der Vergleichsbedingung normaler Alarm. Der Reaktionszeitunterschied zwischen der Unsicherheitsbedingung und der Bedingung normaler Alarm liegt in Studie 3 mit 0,93 Sekunden deutlich über denen in Studie 2 mit 0,341 Sekunden in der Bedingung 4 Rückmeldesituationen pro Fahrt und 0,672 Sekunden bei 12 Rückmeldesituationen pro Fahrt. Diese Beobachtungen stützen die Hypothese zu Ziel 1.

Diese Ergebnisse weisen auf die Effektivität der Kontrollmaßnahme „Variation der Rückmeldesituation“ hin und auch darauf, dass in Studie 2 tatsächlich eine Konfundierung des Cry Wolf Effektes stattgefunden hat. Jedoch sind die Ergebnisse aus Studie 2 und Studie 3 nicht eindeutig vergleichbar, da neben der Variation der Situation, weitere Unterschiede der Studien vorlagen. Anders als beim Vergleich dieser beiden Rückmeldungsarten in Studie 2, wurden in Studie 3 die Anzahl der Rückmeldungssituationen ohne folgenden Automationsfehler (Rückmeldungshäufigkeit) nicht variiert. In Studie 2 konnte zwischen den beiden Häufigkeitsbedingungen kein eindeutiger Unterschied festgestellt werden. Das Verhältnis von Situationen mit und ohne Rückmeldung wurde in Studie 3 so gewählt, dass es in etwa zwischen den beiden Häufigkeitsbedingungen aus Studie 2 liegt. Auch liegt der gefundene mittlere Reaktionszeitunterschied in Studie 3 mit 0,93 Sekunden über dem der beiden Häufigkeitsbedingungen aus Studie 2 (je 0,341 und 0,672 Sekunden). Somit ist es unwahrscheinlich, dass der deutlichere Unterschied zwischen Unsicherheit und normalem Alarm in Studie 3 nicht auf ein deutlich anderes Verhältnis aus Rückmeldungs- und klaren Situationen zurückgeführt werden kann.

Bei den abhängigen Variablen „Nebenaufgabenbearbeitung“ und „situitives Vertrauen“ zeigt sich, wenn überhaupt, nur ein augenscheinlicher aber kein signifikanter Unterschied zwischen der Unsicherheits- und der normaler Alarm-Bedingung. Da vermutet wurde, dass sich ein Cry Wolf Effekt in Rückmeldesituationen auch in einer stärkeren Abwendung zu einer Zweitaufgabe äußert, wurde eigentlich erwartet, dass sich die Unterschiede zwischen den beiden Rückmeldungsarten, zumindest auch auf Ebene der Nebenaufgabenleistung zeigen

würden. Die Nebenaufgabenleistung unterscheidet sich aber, weder insgesamt noch in ihrer Entwicklung über die Fahrt, systematisch zwischen der Unsicherheitsrückmeldung und dem normalen Alarm. Auch bei einer Unsicherheitsrückmeldung wächst die Nebenaufgabenleistung über die Fahrt an. Diese Beobachtung wurde auch schon in den vorherigen Studien gemacht. Die wiederholte Beobachtung und der Fakt, dass es einen deutlichen Unterschied auf Ebene der Reaktionszeit gibt, kann darauf zurückzuführen sein, dass in den unterschiedlichen Rückmeldungsbedingungen unterschiedliche Bearbeitungsstrategien verwendet werden. Solche unterschiedlichen Bearbeitungsstrategien haben zwar eine ähnliche Nebenaufgabenleistung zur Folge, führen aber zu einer unterschiedlichen Reaktionsfähigkeit bei einem Automationsfehler.

Die Verwendung eines Eyetracking-Systems könnte hier genauere Rückschlüsse zulassen. So kann es sein, dass z.B. in der Unsicherheitsbedingung häufiger kurz zur Nebenaufgabe geschaut wird und in der Bedingung normaler Alarm dafür länger am Stück.

Es kann aber auch sein, dass die Nebenaufgabenleistung kein gutes Maß für die Quantifizierung eines Cry Wolf Effektes darstellt. Insbesondere mit der Aussicht auf eine Belohnung bei Erreichung einer möglichst hohen Nebenaufgabenleistung, kann die Nebenaufgabenbearbeitung durch die Versuchspersonen so priorisiert worden sein, dass sie zur Feststellung von Unterschieden im Cry Wolf Effekt insensitive wurde.

So kann es sein, dass es in der Unsicherheitsbedingung zwar eine recht hohe Zuwendung zu der Nebenaufgabe gibt, die Reaktionsbereitschaft jedoch, durch eine höhere Unsicherheit bezüglich des Eintretens eines möglichen Automationsfehlers, höher ist als in der Bedingung normaler Alarm. Trotz einer vergleichbaren Nebenaufgabenleistung kann so eine deutlich bessere Reaktionszeit auf den Automationsfehler erreicht werden.

Auch bei dem situativen Vertrauen zeigt sich kein signifikanter Unterschied.

Deskriptiv liegt das situative Vertrauen in der Unsicherheitsbedingung zwar über dem situativen Vertrauen des normalen Alarms, jedoch ist der Unterschied insgesamt unsystematisch. Es ist möglich, dass das situative Vertrauen weniger durch das Erleben der anscheinend unzuverlässigen Rückmeldung, als vielmehr durch die Gesamtleistung der Automation beeinflusst wird. So kann das Vertrauen auch dadurch positiv beeinflusst werden, dass eine Automation, trotz warnender Rückmeldung, die Fahraufgabe zur Zufriedenheit des Fahrers löst.

Aus dem beobachteten deutlichen Unterschied der Kontrollierbarkeitsleistung zwischen der Unsicherheitsrückmeldung und dem normalen Alarm lässt sich schlussfolgern, dass eine Unsicherheitsrückmeldung nicht einfach nur eine andere

Ausgestaltung eines normalen Alarms ist, sondern auf andere Wirkmechanismen zurückgreift. Dies stützt die in Abschnitt Leithypothesen postulierte These, dass eine Unsicherheitsrückmeldung, im Gegensatz zu einem normalen Alarm, keine sichere Zusammenhangshypothese zwischen dem Rückmeldeereignis und einer negativen Konsequenz erzeugt. Diese Aussage gilt zumindest für den Fall, dass die Erfahrung eines Automationsfehlers nicht bereits vorher schon gemacht wurde. Effekte einer solchen Erfahrung werden im Diskussionsabschnitt zu Ziel 3 diskutiert.

**Ziel 2:** Bestimmung der absoluten Größe eines Cry Wolf Effektes bei Unsicherheitsrückmeldung und normalem Alarm durch einen Vergleich mit einer Bedingung idealer Alarm.

Zur Bewertung eines möglichen Vorteils von unsicherheitsbasierten Rückmeldungen gegenüber normalen Alarmen, gehört nicht nur die Betrachtung der Differenz zwischen diesen beiden Rückmeldungsarten, also die relative Cry-Wolf-Anfälligkeit, sondern auch die Bestimmung der absoluten Cry Wolf Anfälligkeit.

Selbst wenn eine Unsicherheitsrückmeldung im Vergleich zu einem normalen Alarm relativ gut abschneidet, so kann der insgesamt vorliegende Cry Wolf Effekt immer noch zu hoch sein. Es wird ferner postuliert, dass eine Unsicherheitsrückmeldung gar nicht gegenüber einem Cry Wolf Effekt anfällig ist. Dazu wurde der Vergleich der Unsicherheitsrückmeldung und des normalen Alarms zu einem idealen Alarm durchgeführt, bei dem kein Cry Wolf Effekt entstehen kann (siehe Methode Studie 3).

Die Ergebnisse des ersten Fahrtabschnitts zeigen, dass sich die Unsicherheitsrückmeldung im Bereich der Reaktionszeit tatsächlich kaum von einem idealen Alarm unterscheidet, ein normaler Alarm unterscheidet sich hingegen deutlich vom idealen Alarm.

Die Ergebnisse der Nebenaufgabenbearbeitung zeigen deskriptiv einen Unterschied zu der Bedingung idealer Alarm. Die Nebenaufgabenleistung beider Rückmeldungsbedingungen liegt unter der Nebenaufgabenleistung der Bedingung idealer Alarm. Der Unterschied ist zwar nicht signifikant, jedoch ist eine gewisse Systematik zu erkennen. Beide Rückmeldungsarten führen demnach dazu, dass in einer Rückmeldesituation weniger Abwendung zu einer Nebenaufgabe stattfindet als in der gleichen Situation ohne Rückmeldung. Die Rückmeldung scheint also, unabhängig von der Art, eine Wirkung zu haben. Über den Verlauf des ersten Fahrtabschnitts gab es einen tendenziell signifikanten Anstieg der Nebenaufgabenleistung, jedoch keine Interaktion aus Rückmeldungsart und Situationsabfolge. Dies lässt den Schluss zu, dass die Nebenaufgabenleistung in allen drei Bedingungen ähnlich leicht ansteigt, dieser Anstieg also unabhängig von

der Rückmeldungsart ist. Bei einem Cry Wolf Effekt wäre eine Interaktion in der Art erwartet worden, dass sich über die Fahrt die Nebenaufgabenleistung der Unsicherheitsrückmeldung und/oder des normalen Alarms, der des idealen Alarms annähert. Im schlimmsten Fall hätte die Nebenaufgabenleistung im späteren Verlauf der Fahrt auf gleichem Niveau wie der in Nicht-Rückmeldungssituationen gelegen. Die mögliche Insensitivität der Nebenaufgabenleistung für die Detektion eines Cry Wolf Effektes wurde bereits im vorherigen Abschnitt diskutiert.

Für den gleichartigen Anstieg der Nebenaufgabenleistung über die Fahrt in allen drei Rückmeldungsbedingungen scheinen zwei sich nicht unbedingt gegenseitig ausschließende Erklärungen plausibel.

Nach der ersten Erklärung ist der Anstieg auf einen Lerneffekt in der Nebenaufgabenbearbeitung zurückzuführen, der nicht dem Cry Wolf Effekt entspricht. Die Versuchspersonen lernen demnach, auch in Situationen mit Rückmeldung, die Aufgabe effizienter zu bearbeiten.

Die zweite Erklärung würde die gesteigerte Nebenaufgabenbearbeitung mit einer erhöhten Hinwendung zu der Nebenaufgabe gleichsetzen, welche wiederum durch ein gesteigertes Vertrauen in die Automation insgesamt begründet ist. Da auch in der Bedingung idealer Alarm eine tendenzielle Steigerung der Nebenaufgabenleistung beobachtet wird, kann es sein, dass die steigende Nebenaufgabenleistung nicht direkt mit einer sinkenden Wirkung der Rückmeldungseffektivität zusammenhängt, sondern entweder mit einem Anstieg in das Vertrauen in die Automation insgesamt, oder mit einfachen Trainingseffekten.

Die Ergebnisse zum situativen Vertrauen zeigen deskriptiv, dass die Unsicherheitsrückmeldung und der ideale Alarm tendenziell eher auf einem Niveau liegen und das situative Vertrauen in der Bedingung normaler Alarm teilweise deutlich darunter. Die Standardabweichungen sind hier jedoch so groß, dass keine signifikanten Unterschiede beobachtet werden können. Mögliche Probleme dieses Maßes wurden bereits im vorherigen Abschnitt diskutiert.

Im zweiten Fahrtabschnitt, also vor dem Hintergrund der Erfahrung einer Rückmeldesituation mit Automationsfehler, fallen diese Unterschiede teilweise deutlich anders aus. Darauf wird im Diskussionsabschnitt zu Ziel 3, dem Vergleich des ersten und zweiten Fahrtabschnitts, genauer eingegangen.

Die Ergebnisse, insbesondere der Reaktionszeit und, zumindest auf deskriptiver Ebene, die Ergebnisse des situativen Vertrauens, weisen darauf hin, dass eine Unsicherheitsrückmeldung und ein idealer Alarm vergleichbar sind. Diese Beobachtung stützt die Hypothese zu Ziel 2. Ferner wird die These gestützt, dass die Unsicherheitsrückmeldung anders interpretiert wird als ein normaler Alarm. Es wird keine gerichtete Zusammenhangshypothese zwischen Rückmeldung und einer erwarteten Konsequenz aufgebaut, ergo kann fortlaufende negative Evidenz keine

Schwächung einer solchen Hypothese darstellen und ein Cry Wolf Effekt kann nicht resultieren.

Vor dem Hintergrund, dass bis zum ersten Automationsfehler keine Rückmeldung mit nachfolgendem Automationsfehler erlebt wurde, ist die Unsicherheitsrückmeldung einem normalen Alarm gegenüber im Vorteil.

**Ziel 3:** Untersuchung rückmeldungsartspezifischer Lerneffekte aufgrund von Evidenzevaluation.

Die Ergebnisse der Betrachtung des zweiten Fahrtabschnitts (Situation 32 – 63), zeigen ein teilweise deutlich anders geartetes Bild als noch im ersten Fahrtabschnitt. Es sollte untersucht werden, wie sich die Erfahrung einer Rückmeldungssituation mit Automationsfehler auf die weitere Entwicklung eines Cry Wolf Effektes auswirkt. Es wurde postuliert, dass die Erfahrung einer Rückmeldesituation mit Automationsfehler von der Art der Rückmeldungsgestaltung abhängig ist. Insbesondere in der Bedingung normaler Alarm wurde postuliert, dass eine gerichtete Zusammenhangshypothese zwischen Rückmeldung und nachfolgendem Automationsfehler besteht, welche im ersten Fahrtabschnitt, durch die wiederholte Erfahrung einer Rückmeldung ohne Automationsfehler abgeschwächt wird. Grund der Abschwächung ist, dass diese Erfahrung eine negative Evidenz der initialen Zusammenhangshypothese darstellt. Ein tatsächlicher Automationsfehler wird entsprechend nicht mehr oder kaum noch erwartet.

Die unerwartete Erfahrung positiver Evidenz in der ersten Situation mit Automationsfehler führt im weiteren Verlauf der Fahrt zu einer hohen Unsicherheit bezüglich des Zusammenhangs zwischen Rückmeldung und Eintreten eines Automationsfehlers. Positive und negative Evidenz stehen in diesem Fall miteinander in Konkurrenz und lassen die Bildung einer eindeutigen Regel bzw. Zusammenhangshypothese nicht zu. Die Bildung eines Cry Wolf Effektes ist im weiteren Verlauf der Fahrt deutlich gehemmt.

Der Vergleich der Ergebnisse des ersten und zweiten Fahrtabschnittes weist deutlich in diese Richtung. Bei der Reaktion auf einen Automationsfehler reagieren im zweiten Fahrtabschnitt nun alle Versuchspersonen der Bedingung normaler Alarm und können erfolgreich eine Kollision vermeiden. Bezüglich der Reaktionszeit zeigt sich eine deutliche und signifikante Verbesserung von durchschnittlich 1,7 Sekunden.

Die Reaktionszeit der Bedingung normaler Alarm liegt im zweiten Fahrtabschnitt auf etwa dem gleichen Niveau der Bedingung Unsicherheitsrückmeldung, welche sich zwar auch von der ersten zur zweiten Automationsfehlersituation um ca. 0,5

Sekunden leicht verbessert, aber auch deutlich hinter der Reaktionszeitverbesserung der Bedingung normaler Alarm zurückbleibt.

Dass die Reaktionszeit in der Bedingung Unsicherheit von der ersten zur zweiten Automationsfehlersituation recht stabil bleibt, aber in der Bedingung normaler Alarm deutlich verbessert wird, wurde vorher postuliert. Die, in Abhängigkeit des Erfahrungsgrades eines Automationsfehlers, differentielle Wirkung der beiden Rückmeldungsbedingungen, konnte durch einen signifikanten Interaktionseffekt der Faktoren Rückmeldungsart und Fehlererfahrung belegt werden.

Die geringe Veränderung der Reaktionszeit in der Bedingung Unsicherheit im Vergleich zur recht großen Veränderung in der Bedingung normaler Alarm, stützt die These, dass durch die Formulierung der Rückmeldung als Unsicherheit, bei den Versuchspersonen initial eine hohe Unsicherheit bezüglich des Zusammenhangs zwischen Rückmeldung und Automationsfehler besteht und eben keine eindeutige Zusammenhangshypothese.

Die erhöhte Effektivität des normalen Alarms beim zweiten Automationsfehler ist aber nur durch die Erfahrung eines nicht bzw. kaum erwarteten Automationsfehlers erreichbar. Die Unsicherheit bezüglich des Zusammenhangs zwischen Rückmeldung und Automationsfehler wird nicht direkt durch die Art der Rückmeldungsgestaltung vermittelt, wie es im Fall der Unsicherheitsrückmeldung ist, sondern erst durch die Erfahrung widersprüchlicher bzw. negativer Evidenz.

Die Reaktionszeitergebnisse der Bedingungen Unsicherheit und normaler Alarm müssen aber auch in Relation zu den Ergebnissen der Bedingung idealer Alarm gesehen werden. In dieser Bedingung liegt die Reaktionszeit beim ersten Automationsfehler auf einem vergleichbaren Niveau mit der Bedingung Unsicherheit, die Reaktionszeit verbessert sich zum zweiten Automationsfehler aber deutlich (Verbesserung um etwa 1,9 Sekunden) und ist dort eindeutig niedriger als in der Bedingung Unsicherheit. Die Verbesserung in der Bedingung idealer Alarm ist damit, relativ gesehen, ähnlich gut wie in der Bedingung normaler Alarm, jedoch liegen die Reaktionszeiten insgesamt deutlich besser als beim normalen Alarm.

Die Reaktionszeitverbesserung in der Bedingung idealer Alarm zeigt auch, dass im ersten Fahrtabschnitt der ideale Alarm nicht maximal effizient war, sondern diese Effizienz erst nach der Erfahrung eines Automationsfehlers erreicht werden konnte. Die nicht optimale Effizienz des idealen Alarms in der ersten Automationsfehlersituation kann nicht auf einen Cry Wolf Effekt zurückzuführen sein, es wurden ja vorher keine falschen Alarme erlebt.

Hier können zwei Erklärungsmöglichkeiten plausibel sein. So ist es möglich, dass sich ein generelles Übervertrauen in die Automation entwickelt, die insgesamt die Erwartung eines Automationsfehlers verringert, bzw. die Schwere eines möglichen

Automationsfehlers so nicht erwarten lässt. Die Entstehung des Übervertrauens kann z.B. dadurch verstärkt werden, dass in der Bedingung idealer Alarm im Verlauf der Fahrt keine Rückmeldung erlebt wird, wohingegen in den Bedingungen normaler Alarm und Unsicherheit das wiederholte Erleben der Rückmeldungen der Entstehung eines allzu großen Übervertrauens entgegenwirkt.

Alternativ kann es aber auch sein, dass im ersten Fahrtabschnitt die Versuchspersonen noch nicht genau wissen, wie die Situation am besten zu kontrollieren ist, bzw. das Handlungsschema „Vollbremsung“ noch nicht voraktiviert ist. In der zweiten Automationsfehlersituation kann diese Handlung dann schneller abrufbar sein. Letztere Erklärung würde aber in gleicher Weise für alle Rückmeldungsarten gelten und wäre damit konstant gehalten.

Der Vergleich der Reaktionszeitdaten der Bedingung idealer Alarm und der Bedingungen Unsicherheit und normaler Alarm zeigt aber auch, dass sowohl in der Bedingung Unsicherheit, als auch in der Bedingung normaler Alarm, selbst mit der gemachten Erfahrung eines Automationsfehlers, eine optimale Reaktionszeit, wie sie bei einem idealen Alarm möglich sein könnte, nicht erreicht wird. Dies gilt zumindest für den eingeschränkten Beobachtungsrahmen von nur zwei erlebten Automationsfehlern. Dies ist wiederum ein Hinweis darauf, dass bei Rückmeldungen mit einer relativ hohen Unsicherheit bezüglich des Eintretens und Art einer Konsequenz, die Erwartung eines Automationsfehlers nicht maximal ist und ein Cry Wolf Effekt nicht vollkommen vermieden werden kann.

Versuchspersonen neigen wohlmöglich dazu ein sogenanntes „Probability Matching“ zu betreiben (Bliss & Dunn, 2000), indem sie ihre Reaktion einer Art subjektiv gebildeten Wahrscheinlichkeit des Eintretens eines tatsächlichen Automationsfehlers folgen lassen. Diese subjektive Wahrscheinlichkeit wäre im Fall eines maximalen Cry Wolf Effektes sehr niedrig, bei einem Idealen Alarm entsprechend sehr hoch und bei Rückmeldungen mit hoher Unsicherheit (Unsicherheitsrückmeldung und normaler Alarm mit erlebter ambivalenter Evidenz) tendenziell in einem mittleren Bereich.

Wenn das Alarmsystem anfällig für falsche Alarme ist, kann durch die Unsicherheitsrückmeldung zumindest bei dem ersten erlebten Automationsfehler eine bessere Reaktion erzeugt werden als bei einem normalen Alarm, jedoch insgesamt keine ideale Reaktion, bzw. wenn, dann nur in Ausnahmefällen.

Die Umsetzung eines idealen Alarmsystems ist, nach Überlegungen der Signalentdeckungstheorie, in einem System mit „verrauschten“ Daten jedoch kaum möglich, die Unsicherheitsrückmeldung stellt daher trotzdem eine gute Alternative dar.

Die Erfahrung eines Cry Wolf Effektes bei einem normalen Alarm in der ersten Fehlersituation führt zwar im zweiten Fahrtabschnitt zu einer höheren Alarmierung und somit besseren Reaktionsfähigkeit, aber auch zu einem deutlich geringeren

situativen Vertrauen in die Automation. Obwohl im zweiten Fahrtabschnitt die Reaktionszeiten vergleichbar sind mit denen der Bedingung Unsicherheit, liegt das situative Vertrauen im zweiten Fahrtabschnitt deskriptiv teilweise deutlich unter dem der Bedingungen Unsicherheit und idealer Alarm, die beide auf einer konstant hohen Ebene liegen. Die paarweisen Vergleiche zeigen zwar keine signifikanten Unterschiede, die Irrtumswahrscheinlichkeiten liegen für die Vergleiche normaler Alarm – Unsicherheit und normaler Alarm – idealer Alarm aber sehr deutlich niedriger als für den Vergleich Unsicherheit – idealer Alarm. Es wird auch ein tendenzieller Interaktionseffekt der Faktoren Rückmeldung und Situationsabfolge beobachtet. Während das situative Vertrauen in den Bedingungen Unsicherheit und idealer Alarm recht konstant bleibt, schwankt das situative Vertrauen in der Bedingung normaler Alarm teils erheblich.

Das niedrigere Vertrauen in die Automation in der Bedingung normaler Alarm kann durch den starken Cry Wolf Effekt beim ersten Automationsfehler erklärt werden. Da hier in der Bedingung normaler Alarm ein deutlicher Cry Wolf Effekt vorliegt, kann der Automationsfehler schlecht kontrolliert werden. Dies führt dazu, dass die Situation als deutlich bedrohlicher wahrgenommen wird als in den Bedingungen Unsicherheit und idealer Alarm. Die antizipierte Konsequenz nach einer Rückmeldung erhält somit einen stark aversiven Charakter. Dies, gepaart mit der hohen Unsicherheit bezüglich des Eintretens eines Fehlers aufgrund der gemachten Erfahrung eines unzuverlässigen Alarms, führt zwar zu einer deutlich verbesserten Reaktionsbereitschaft, aber auch zu einem deutlich niedrigeren situativen Vertrauen in die Automation. Dieses Ergebnis ist ein weiterer deutlicher Hinweis, dass in der ersten Automationsfehlersituation in der Bedingung normaler Alarm ein deutlich ausgeprägter Cry Wolf Effekt stattfand als in den Vergleichsbedingungen. Dies stützt die Hypothese, dass eine Unsicherheitsrückmeldung nicht mit einer gerichteten Zusammenhangshypothese verbunden ist und Rückmeldungen ohne Automationsfehler nicht als negative Evidenzen zu einer solchen Zusammenhangshypothese betrachtet werden.

Bei der abhängigen Variable Nebenaufgabenleistung, wird die Nebenaufgabenleistung zwischen dem ersten und zweiten Fahrtabschnitt miteinander verglichen. Vermutet wurde, dass es, nach dem Erleben des ersten Automationsfehlers, in der Bedingung normaler Alarm im zweiten Fahrtabschnitt zu einer niedrigeren Nebenaufgabenleistung kommt. In den Bedingungen Unsicherheit und idealer Alarm hingegen wurde vermutet, dass die Nebenaufgabenleistung im zweiten Fahrtabschnitt gleichbleibt. Es zeigt sich jedoch, dass die Nebenaufgabenleistung in der Bedingung normaler Alarm auf einem gleichen Niveau bleibt, bzw. kaum sinkt, die Nebenaufgabenleistung in den Bedingungen



Unsicherheit und idealer Alarm im zweiten Fahrtabschnitt dafür im Durchschnitt deutlich höher ist. Es gibt eine signifikante Interaktion aus der Rückmeldungsart und der Reihenfolge des Automationsfehlers. Da die Nebenaufgabenleistung im ersten Fahrtabschnitt zwischen normalem Alarm und Unsicherheit noch vergleichbar ist, kann davon ausgegangen werden, dass das Erleben des Automationsfehlers in der Bedingung normaler Alarm ein weiteres Ansteigen der Nebenaufgabenleistung hemmt. In der Bedingung Unsicherheit scheint das Erleben des Automationsfehlers jedoch deutlich weniger Auswirkungen auf die weitere Entwicklung der Nebenaufgabenleistung zu haben.

Die Beobachtung entspricht zwar nicht exakt der Hypothese, weist aber sinngemäß eine ähnliche Tendenz auf. Der Automationsfehler am Ende des ersten Fahrtabschnitts scheint demnach in der Bedingung Unsicherheit eher erwartet zu werden und hat daher einen weniger aversiven Charakter. In der Bedingung normaler Alarm scheint hier die Erwartung deutlich geringer zu sein, was den aversiven Charakter der Situation erhöht und zu einer deutlichen Unsicherheit in weiteren Rückmeldesituationen führt. Insgesamt scheint dies zu einer Verunsicherung der Versuchspersonen zu führen. Diese Verunsicherung hält sie von einer weiteren Optimierung der Nebenaufgabenbearbeitung, wie sie in den Bedingungen Unsicherheit und idealer Alarm stattfindet, ab. Auch dieses Ergebnis stützt tendenziell die Hypothese, dass Unsicherheit und normale Alarmer unterschiedlich interpretiert werden.

**Ziel 4:** Untersuchung der Auswirkung einer gemischten Alarmstrategie aus Unsicherheitsrückmeldung und normalem Alarm auf die Effektivität der Rückmeldungsstufe Unsicherheit.

In der Längsschnittbetrachtung über die zwei Fahrsituationen wird auch eine gemischte Rückmeldestrategie untersucht, welche in der ersten Automationsfehlersituation einen normalen Alarm rückmeldet und in der zweiten Automationsfehlersituation Unsicherheit. In den unklaren Fahrsituationen ohne Automationsfehler wird im Verlauf der gesamten Fahrt immer Unsicherheit zurückgemeldet.

Solche Mischstrategien, wie sie z.B. Likelihoodalarme darstellen (Sorkin, Kantowitz & Kantowitz, 1988), könnten z.B. verwendet werden, wenn feststellbar ist, dass ein Automationsfehler in einer Situation wahrscheinlicher ist als in einer anderen Situation. Durch eine von dieser unterschiedlichen Wahrscheinlichkeit abhängige unterschiedliche Rückmeldungsart kann die Dringlichkeit des Eingreifens unterstrichen werden.

Es ist jedoch auch möglich, dass in den Situationen mit geringerer Wahrscheinlichkeit eines Automationsfehlers, ein solcher trotzdem passiert. Die Effektivität der Alarmstufe darf also nicht auf Kosten der Likelihoodstufe gehen. Es wurde postuliert, dass das Mischen unterschiedlicher Rückmeldungen zu einer Schwächung der Effektivität einer der Rückmeldungen führt, da eine Erwartungshypothese aufgebaut wird, die einen Zusammenhang zwischen Automationsfehlern und der einen Art der Rückmeldung (in diesem Fall der normale Alarm) und keinen Zusammenhang zwischen Automationsfehlern und der anderen Rückmeldungsart (in diesem Fall Unsicherheit) postuliert. Entsprechend müsste es in der ersten Automationsfehlersituation eine Reaktion ähnlich gut wie in der Bedingung idealer Alarm geben, da der normale Alarm hier zum ersten Mal geäußert wird und entsprechend kein Cry Wolf Effekt entstanden sein kann. In der zweiten Automationsfehlersituation müsste die Reaktionszeit deutlich höher sein als in der Vergleichsbedingung Unsicherheitsrückmeldung.

Der Vergleich der Reaktionszeit der ersten und zweiten Automationsfehlersituation in den Bedingungen Unsicherheit, idealer Alarm und Mischstrategie zeigt entsprechend in der ersten Automationsfehlersituation keinen Unterschied zwischen den drei Rückmeldungsbedingungen. In der zweiten Automationsfehlersituation ist die Reaktionszeit, bei deskriptiver Betrachtung, in der Bedingung Unsicherheit gegenüber der ersten Situation verbessert, in der Bedingung Mischstrategie hingegen auf dem gleichen Niveau der ersten Situation geblieben. Es konnte jedoch kein signifikanter Interaktionseffekt festgestellt werden.

Dieses Ergebnis weist darauf hin, dass die Unsicherheitsrückmeldung deutlich robust gegen einen Cry Wolf Effekt ist. In der Bedingung Mischstrategie wurde nämlich erst in der zweiten Automationsfehlersituation das erste Mal eine Kopplung aus Unsicherheitsrückmeldung und Automationsfehler erlebt und somit vorher erheblich mehr Rückmeldungssituationen ohne Automationsfehler als vor der ersten Automationsfehlersituation. Die Reaktionszeit bleibt dennoch auf dem gleichen Niveau der ersten Automationsfehlersituation. Es entsteht also kein weiterer oder höher ausgeprägter Cry Wolf Effekt, obwohl deutlich mehr „falsche Alarmer“ erlebt wurden.

Dennoch weisen die Ergebnisse in die vermutete Richtung, dass das Mischen der Rückmeldestrategien, zu einer Minderung der Effektivität einer der Rückmeldungsarten führt. In der Bedingung Unsicherheitsrückmeldung findet nämlich eine Verbesserung der Reaktionszeit statt, die in der Bedingung Mischstrategie ausbleibt. Die aktuelle Datenlage bietet jedoch keine hinreichende Evidenz. Um hier eine eindeutige Aussage treffen zu können, sind weitere Untersuchungen mit mehr Versuchspersonen erforderlich.

**Ziel 5:** Untersuchung der Auswirkung einer Unsicherheitsrückmeldung auf die Erwartung eines Automationsfehlers in Situationen ohne Rückmeldung (Reliance).

Bei der Betrachtung der Auswirkung unterschiedlicher Rückmeldearten auf die Kontrollierbarkeit eines Automationsfehlers in sogenannten Miss-Situationen, sollte insbesondere der Einfluss der Rückmeldungsart auf die Reliance in das Rückmeldesystem untersucht werden, also der Grad der Überzeugung, dass im Falle eines berechtigten Grundes für einen Alarm, das Alarm bzw. Rückmeldesystem anschlägt. Es wurde angenommen, dass es eine Abhängigkeit von Compliance und Reliance gibt. Es wurde ferner angenommen, dass der Grad der Compliance, worunter auch der Cry Wolf Effekt fällt, von der Art der Rückmeldung abhängt. In diesem Fall könnte sich die Art der Rückmeldung und der mit dieser Rückmeldung gemachte Erfahrungskontext auch auf die Ausprägung der Reliance auswirken. Aus ökonomischen Gründen konnte, im Rahmen von Studie 3, dieser Vergleich jedoch nicht mehr in der Art kontrolliert und systematisch variiert werden, wie es notwendig gewesen wäre. Die Betrachtung der Miss Situationen wird daher eher als exploratorisch angesehen und es können lediglich Tendenzen aufgezeigt werden. Eine genauere Untersuchung sollte in weiterführenden Arbeiten durchgeführt werden.

Konkret sollte die Hypothese überprüft werden, dass ein Cry Wolf Effekt (niedrige Compliance in ein Alarmsystem) ebenso zu einer niedrigeren Reliance in das Alarmsystem führen würde. Da vermutet wurde, dass es insbesondere bei einem normalen Alarm zu einem Cry Wolf Effekt kommt, wird erwartet, dass dies wiederum in Miss Situationen ein Vorteil gegenüber der Unsicherheitsrückmeldung darstellen könnte. Im Vergleich dazu hätte der Ideale Alarm zu maximaler Compliance und demzufolge auch zu einer maximalen Reliance führen müssen, also einer besonders schlechten Reaktion in einer Miss-Situation.

Die Ergebnisse des Reaktionszeitvergleichs der Miss-Situation zeigen jedoch nur sehr geringe und nicht signifikante Unterschiede zwischen dem normalen Alarm und der Unsicherheitsrückmeldung. Bei den gegebenen Irrtumswahrscheinlichkeiten ist von einer annähernden Äquivalenz der Reaktionszeiten auszugehen ( $p = .76$ ). Die Reaktionszeit in der Miss-Situation liegt in allen drei Rückmeldungsbedingungen (Unsicherheit, normaler Alarm und idealer Alarm) signifikant über den Reaktionszeiten der zweiten Automationsfehlersituation und ungefähr auf dem Niveau der Reaktionszeit der Bedingung normaler Alarm in der ersten Automationsfehlersituation.

Diese Beobachtung spricht zunächst gegen die Vermutung einer Abhängigkeit von Compliance und Reliance und wäre konform mit den Beobachtungen von Dixon, Wickens & McCarley(2007). Grundsätzlich ist es möglich, dass das Vorhandensein

eines Rückmeldesystems, unabhängig von der Art der Rückmeldung, eine hohe Reliance erzeugt. Wenn ein Rückmeldesystem vorhanden ist, dann wird anscheinend vorausgesetzt, dass dies einen Fehlerfall abdeckt.

Einschränkend bei der Betrachtung der Miss-Situation ist der Umstand, dass im vorherigen Verlauf der Fahrt bereits zwei Situationen mit einem Automationsfehler erlebt wurden. Ein möglicher Cry Wolf Effekt kann damit in seiner Ausprägung schon stark verringert gewesen sein. In der Bedingung normaler Alarm spricht die in der zweiten Automationsfehlersituation deutlich besseren Reaktionszeit als in der ersten Automationsfehlersituation dafür, dass ein Automationsfehler zumindest implizit erwartet wurde. Dies spricht gegen eine niedrige Compliance oder zumindest für eine ähnlich hohe Fehlererwartung, sowohl in der Unsicherheitsbedingung als auch in der Bedingung normaler Alarm.

Für eine systematische Untersuchung der Abhängigkeit von Compliance und Reliance hätte die Compliance in die beiden Alarmsysteme aber unterschiedlich hoch sein sollen, wie es in der ersten Automationsfehlersituation wahrscheinlich noch der Fall gewesen ist.

Die Reaktionszeit in der Miss-Situation liegt in allen drei Rückmeldungsbedingungen (Unsicherheit, normaler Alarm und idealer Alarm) signifikant über den Reaktionszeiten der zweiten Automationsfehlersituation und ungefähr auf dem Niveau der Reaktionszeit der Bedingung normaler Alarm in der ersten Automationsfehlersituation. Dies spricht dafür, dass die Reaktionszeit der Bedingung normaler Alarm in der ersten Automationsfehlersituation bereits auf dem Niveau von keinerlei Fehlererwartung gewesen ist.

Konsequenterweise hätte ein Vergleich zu einer Bedingung vollkommen ohne Rückmeldung durchgeführt werden müssen. Dies konnte aber im Rahmen der Studie 3 aus ökonomischen Gründen nicht geschehen.

Eine ungefähre Abschätzung erlaubt der Vergleich zu Reaktionszeitwerten der Bedingung keine Rückmeldung aus Studie 1. Eine genaue Vergleichbarkeit ist zwar nicht gegeben, da die Anzahl der erlebten Rückmeldungen ohne nachfolgenden Automationsfehler variiert. Jedoch ist die eigentliche Kontrollierbarkeitssituation vom Erscheinungsbild und zeitlichem Ablauf her exakt zwischen Studie 1 und 3 vergleichbar. In Studie 1 konnte in der Bedingung „Keine Rückmeldung“ eine Reaktionszeit von etwa 5 Sekunden beobachtet werden (sowohl bei vielen als auch wenigen vorher erlebten Rückmeldungen ohne Automationsfehler). Die Reaktionszeit der Miss-Situationen liegt in allen drei Rückmeldungsbedingungen auch bei etwa 5 Sekunden. Dies lässt vermuten, dass die Reaktionszeit in der Miss Situation tatsächlich auf dem Niveau von keiner Fehlererwartung liegt. Die Reaktionszeit in Miss Situationen ist allerdings auch nicht schlechter als in einer Bedingung ohne Rückmeldung.

Bezüglich der Auswirkungen des Erlebens einer Miss Situation auf das Vertrauen in das Alarmsystem, wird das Gesamtvertrauen in die Automation nach der Miss-Situation mit dem Gesamtvertrauen in die Automation nach dem zweiten Automationsfehler verglichen. Hier kann ein deutlich negativer Effekt des Erlebens einer Miss-Situation auf das Gesamtvertrauen in die Automation beobachtet werden. Unabhängig von der Art der Rückmeldung wirkt sich das Erleben einer Miss Situation gleichermaßen negativ auf das Gesamtvertrauen aus. Dies spricht dafür, dass unabhängig von der Rückmeldungsart ein ähnlicher Grad von Enttäuschung in das Automations- und damit auch das Rückmeldesystem vorliegt.

## **10 Gesamtdiskussion**

Abschließend werden die Ergebnisse der durchgeführten Studien hinsichtlich der eingangs formulierten Leithypothesen diskutiert. Die Leithypothesen sind nach zwei möglichen Effekten aufgeteilt. Zum einen die Effekte der Unsicherheitsrückmeldung hinsichtlich der situativen Erwartung von Automationsfehlern und entsprechend der Kontrollierbarkeit, zum anderen die Effekte hinsichtlich der generellen Erwartung von Automationsfehlern, z.B. durch angepasstes Vertrauen.

### *10.1 Auswirkungen auf die situative Erwartung von Automationsfehlern*

Die **erste Leithypothese** behauptet, dass eine Unsicherheitsrückmeldung grundsätzlich effektiv ist, eine situative Erwartung eines Automationsfehlers zu bilden.

In Studie 1 wurde diese Hypothese, welche als notwendiges Kriterium für die weitere Untersuchung der Unsicherheit als Rückmeldestrategie gilt, durch einen Vergleich zwischen einer Unsicherheitsrückmeldung mit einer Kontrollbedingung ohne Rückmeldung vor Automationsfehlern überprüft. Die Ergebnisse zeigen, dass eine Unsicherheitsrückmeldung grundsätzlich effektiv ist, da sie zu besseren Kontrollierbarkeitsreaktionen bei einem Automationsfehler führt als in der Kontrollbedingung.

In Studie 1 zeigte sich aber auch, dass die Kontrollierbarkeit eines Automationsfehlers in der Unsicherheitsbedingung in Abhängigkeit der erlebten Rückmeldungen ohne folgenden Automationsfehler schlechter wurde. Es stellt sich die Frage, wie groß dieser offensichtliche Cry Wolf Effekt insbesondere im Vergleich zu anderen Alarmstrategien ist.

Diesen möglichen differentiellen Effekt in Abhängigkeit der Rückmeldestrategie adressieren die Leithypothesen 2 – 4.

Die **zweite Leithypothese** gilt als zweites notwendiges Kriterium der Untersuchung der Unsicherheit. Sie besagt, dass ein wiederholtes Erleben des Ausbleibens einer negativen Konsequenz (z.B. ein Automationsfehler) nach einer Rückmeldung einen Cry Wolf Effekt erzeugt. Die Stärke des Cry Wolf Effektes ist dabei proportional zur Menge, der ohne Unterbrechung erlebten Rückmeldungen ohne folgenden Automationsfehler.

In Studie 1 wurde ein Automationsfehler zu zwei unterschiedlichen Zeitpunkten erlebt. Im ersten Fall wurden vorher wenige Rückmeldungen ohne nachfolgenden Automationsfehler erlebt, im zweiten Fall deutlich mehr. Obwohl kein signifikanter Effekt beobachtet werden konnte, zeigte sich deskriptiv ein augenscheinlicher Unterschied. Im Durchschnitt lag die Reaktionszeit höher, wenn Versuchspersonen mehr „falsche Alarme“ erlebt hatten, verglichen mit einer geringen Anzahl vorher erlebter „falscher Alarme“. Dass ein möglicher Effekt hier nicht aufgedeckt werden konnte kann mehrere Gründe haben. Zum einen wurde in Studie 1 nur eine Rückmeldungsart, nämlich die Unsicherheitsrückmeldung, untersucht. Für diese wurde insgesamt eine geringe Cry Wolf Neigung postuliert. Zum anderen scheinen deutliche interindividuelle Unterschiede der Versuchspersonen bezüglich der Anfälligkeit für einen Cry Wolf Effekt eine Rolle zu spielen. Hier gab es große Unterschiede zwischen den Versuchspersonen. Was sich bei der geringen Stichprobengröße in einer recht hohen Varianz der Kontrollierbarkeit des Automationsfehlers äußerte.

In Studie 2 wurde die Anzahl der erlebten Rückmeldungen ohne Automationsfehler vor einem tatsächlichen Fehlerereignis systematisch variiert und zwischen einer Unsicherheitsrückmeldung und einem normalen Alarm verglichen. Es zeigte sich wieder auf deskriptiver Ebene ein Reaktionszeitunterschied zwischen dem Automationsfehler nach 4 falschen Alarmen und dem Automationsfehler nach 12 falschen Alarmen. Dieser Unterschied zeigt sich nur in der Bedingung normaler Alarm, nicht aber in der Bedingung Unsicherheit. Dies steht zunächst im Gegensatz zu den Beobachtungen aus Studie 1, in der bei Unsicherheitsrückmeldung deskriptiv ein Reaktionszeitunterschied vorlag. Der beobachtete Unterschied ist aber weder für die Anzahl der falschen Alarme noch als Interaktion aus Rückmeldungsart und falschen Alarmen signifikant.

Eine Erklärungsmöglichkeit liefert die bereits diskutierte Sekundärvarianz. Zieht man die beobachtete Kollisionshäufigkeit mit ein, so weisen die Beobachtungen

insgesamt in die Richtung der Hypothese. Es kollidieren nämlich deutlich mehr Versuchspersonen, wenn vor dem Automationsfehler 12 Rückmeldungen als „falsche Alarmer“ erlebt wurden.

Daher wird, obwohl sowohl in Studie 1 als auch in Studie 2 keine eindeutigen Belege für einen Effekt gefunden werden konnten, angenommen, dass ein Cry Wolf Effekt in dieser simulierten Umgebung im Kontext des (teil)automatisierten Fahrens erzeugt werden konnte.

Die Stärke des Cry Wolf Effektes scheint abhängig von der Menge, der vorher in einer Reihe erlebten Rückmeldungen ohne Automationsfehler. Diese Aussage gilt zumindest für recht kurze Zeiteinheiten.

Die Menge der erlebten Rückmeldesituationen wurde nämlich in der recht kurzen jeweiligen Versuchsdauer untergebracht. Es lag also eine sehr hohe Dichte dieser Situationen vor. Die Ergebnisse der Studien ermöglichen vor diesem Hintergrund keine Aussagen dazu, in wie weit die Zeit, welche zwischen den einzelnen Rückmeldeereignissen ohne Automationsfehler liegt, einen Einfluss auf die Stärke eines Cry Wolf Effektes haben kann. So ist es möglich, dass nicht nur die Menge der erlebten Rückmeldesituationen, sondern auch die Zeit dazwischen, die Stärke eines Cry Wolf Effektes beeinflussen. Erstrecken sich unterschiedlich häufige Situationen von Rückmeldungen auf einen insgesamt langen Zeitraum, kann es sein, dass dieser Zeitraum die Wirkung der Dichte der Situationen relativiert.

**Leithypothese 3** besagt, dass über die Art der Rückmeldungsgestaltung, Einfluss auf die Cry Wolf Anfälligkeit der Rückmeldung genommen werden kann. Als Erklärung für die Resistenz einer Unsicherheitsrückmeldung gegenüber einem Cry Wolf Effekt wird die Theorie der Evidenzevaluation genutzt.

Der postulierte Einfluss besteht darin, dass unterschiedliche Rückmeldungsgestaltung unterschiedliche Erwartungen bezüglich des Auftretens einer Konsequenz nach einem Alarmereignis erzeugen. Diese Erwartungen resultieren aus unterschiedlichen Hypothesen über den Zusammenhang zwischen Alarmereignis und Konsequenz.

Es wird postuliert, dass die Zusammenhangshypothese eines konventionellen Alarmes das Ausbleiben einer Konsequenz nach einem Alarm nicht vorsieht – so gesehen also „starr“ ist. Das Ausbleiben eines Automationsfehlers nach einem normalen Alarm stellt somit immer negative Evidenz dar.

Eine Unsicherheitsrückmeldung sieht hingegen das Ausbleiben einer Konsequenz nach einem Alarmereignis im Konzept vor. Das Ausbleiben der Konsequenz stellt also eine valide Option nach einer Rückmeldung dar und ist nicht als negative

Evidenz für die Zusammenhangshypothese zu verstehen. Die Zusammenhangshypothese kann daher als flexibel bezeichnet werden.

Daraus wird gefolgert, dass es zwischen einem konventionellen Alarm und einer Unsicherheitsrückmeldung einen Unterschied in dem Sinne geben muss, dass die Unsicherheitsrückmeldung weniger anfällig für die Entwicklung eines Cry Wolf Effektes ist.

Ein Cry Wolf Effekt wurde als geringere Erwartung eines Automationsfehlers nach einem Alarm bzw. Rückmeldeereignis definiert. Die Erwartung wurde anhand unterschiedlicher abhängiger Variablen operationalisiert. Allen voran über die Kontrollierbarkeit einer Automationsfehlersituation (Kollision ja/nein und Reaktionszeit). Des Weiteren die Aufmerksamkeitsverteilung zwischen Fahr und Nebenaufgabe anhand der Nebenaufgabenleistung in den Rückmeldesituationen. Schließlich das subjektiv empfundene situative Vertrauen in die Automation, erfasst kurz nach den Rückmeldesituationen.

Die Ergebnisse aus den Studien 1 & 2 weisen zunächst nicht in die erwartete Richtung, was aber aufgrund des Einflusses von Sekundärvarianz zu erklären ist (siehe Diskussion Studie 2 & 3). In Studie 3, nach der Kontrolle möglicher Sekundärvarianz, konnte schließlich gezeigt werden, dass grundsätzlich ein Unterschied in der Cry Wolf Neigung in Abhängigkeit der Rückmeldungsgestaltung besteht.

Der erste Hinweis zugunsten der Hypothese zeigt sich durch den beobachteten Unterschied zwischen einem konventionellen Alarm und der Unsicherheitsrückmeldung in der Kontrollierbarkeit eines Automationsfehlers, wenn vorher eine Reihe von Rückmeldungen ohne Automationsfehler erlebt wurde. Jedoch zeigte sich auch bei einer Unsicherheitsrückmeldung eine Verschlechterung der Kontrollierbarkeit in Abhängigkeit der Anzahl der in einer Reihe erfolgten Situationen in denen zwar Unsicherheit zurückgemeldet wurde, ein Automationsfehler aber ausblieb. Die Hypothese trifft also vermeintlich nicht vollständig zu, da diese behauptet, dass bei einer Unsicherheitsrückmeldung das Ausbleiben von Automationsfehlern keine negative Evidenz darstellt. Eine mögliche Erklärung ist, dass häufiges Erleben des Ausbleibens eher eine Verlagerung der subjektiv wahrgenommenen Wahrscheinlichkeit eines Automationsfehlers nach einer Rückmeldung bewirkt. Eine Zusammenhangshypothese der Unsicherheitsrückmeldung postuliert ja, der Theorie nach, nur die Möglichkeit des Auftretens eines Automationsfehlers im Sinne einer Wahrscheinlichkeit. Wobei initial keinerlei eindeutige Aussage über die Höhe dieser Wahrscheinlichkeit möglich ist. Das häufige Erleben ausbleibender Konsequenzen scheint diese Wahrscheinlichkeit zu verringern. Die Beobachtung ausbleibender Automationsfehler fällt somit deutlich weniger ins Gewicht als bei einem konventionellen Alarm, da hier, ausgehend von

184



einem starren Zusammenhang, ausbleibende Automationsfehler deutlich diesem Zusammenhang widersprechen.

Eine weitere, sich nicht gegenseitig ausschließende, Erklärung für eine Abnahme der Rückmeldungseffektivität der Unsicherheitsrückmeldung ist die bereits erwähnte Sekundärvarianz aufgrund der hohen Ähnlichkeit der erlebten Situationen (siehe Diskussion Studie 2 & 3). Auch wenn diese in Studie 3 anscheinend zufriedenstellend minimiert wurde, so konnte diese aufgrund der doch relativ einfachen Simulationsumgebung wahrscheinlich nicht eliminiert werden.

Eine weitere Methode der Untersuchung der These, dass das Ausbleiben von Konsequenzen nach einer Rückmeldung bei Unsicherheit keine negative Evidenz darstellt, stellt der Vergleich der Höhe eines Cry Wolf Effektes in Abhängigkeit der Menge der erlebten Evidenzen dar.

Handelt es sich beim Ausbleiben einer Konsequenz nach einer Rückmeldung um negative Evidenz, dann müsste die Zusammenhangshypothese umso schwächer sein, je mehr Rückmeldungen ohne Automationsfehler erlebt wurden. D.h., bei weniger negativer Evidenz kommt es zu einem gering ausgeprägten Cry Wolf Effekt, bei mehr erlebter negativer Evidenz zu einem stärkeren Cry Wolf Effekt etc. Bei einer Unsicherheitsrückmeldung dürfte diese Beziehung entweder gar nicht, oder deutlich geringer ausgeprägt sein (siehe auch Leithypothese 2). Diese Hypothese wurde in Studie 2 adressiert, jedoch kam es hier, aufgrund der bereits erwähnten Sekundärvarianz, zu uneindeutigen Ergebnissen. Werden nur die Fahrer betrachtet, die auch eine Reaktion zeigten – so zeigte sich rein deskriptiv die postulierte Interaktion aus Anzahl potenzieller negativer Evidenz und der Rückmeldungsart. Diese Interaktion ließ sich aber, aufgrund der geringen Größe der verbliebenen Stichprobe, nicht statistisch absichern.

In Studie 3 wurden die beiden Rückmeldestrategien normaler Alarm und Unsicherheit zusätzlich noch zu einem idealen Alarm verglichen, eine Bedingung in der tatsächlich nur dann eine Rückmeldung geäußert wird, wenn es tatsächlich zu einem Automationsfehler kommt. In dieser Bedingung wird also überhaupt keine negative Evidenz erlebt. Da es in der Theorie bei einem idealen Alarm nicht zu einem Cry Wolf Effekt kommen kann, kann die Größe eines Cry Wolf Effektes anhand des Vergleichs zu einem idealen Alarm quantifiziert werden.

In Studie 3 konnte eine vergleichbar gute Qualität der Kontrollübernahme bei idealem Alarm und Unsicherheitsrückmeldung beobachtet werden, wenn ein Automationsfehler das erste Mal erlebt wird. Dies stand im Gegensatz zur Kontrollleistung bei normalen Alarmen, die hier deutlich schlechter ausfielen. Diese Aussage trifft allerdings nur dann zu, wenn ein Automationsfehler das erste Mal erlebt wird. Mit den Auswirkungen der Erfahrung positiver Evidenz auf die Cry Wolf Neigung im weiteren Fahrtverlauf beschäftigt sich Leithypothese 4.

**Leithypothese 4** besagt, dass das Erleben positiver Evidenz eine Zusammenhangshypothese stabilisieren kann. So wird nach dem Erleben eines berechtigten Alarms die Entstehung eines Cry Wolf Effektes zumindest in einer bestimmten Zeit gehemmt.

Wenn bei einer Unsicherheitsrückmeldung das Erleben von Automationsfehlern nicht in der Art positive Evidenz darstellt, wie es bei einem konventionellen Alarm vermutet wird, dann muss es eine Interaktion aus der Anzahl der tatsächlich erlebten Fehler und der Rückmeldungsart geben.

Während es bei einer Unsicherheitsrückmeldung zu keiner bzw. nur geringen Veränderung der Rückmeldungseffektivität eines Automationsfehlers von einem Fehler zum anderen kommt, findet in der Bedingung normaler Alarm zwischen der ersten zur zweiten erlebten Fehlersituation eine deutliche Erhöhung der Rückmeldungseffektivität statt. Die Erkenntnis, dass auf eine Rückmeldung tatsächlich ein Automationsfehler mit möglichen schwerwiegenden Konsequenzen folgen kann, hemmt deutlich die Entstehung eines weiteren Cry Wolf Effektes. Dies entspricht im Wesentlichen den Beobachtungen die auch Bahner (2008) bei der Untersuchung der Einflüsse von Fehlererfahrung auf die Erwartung von Automationsfehlern machte.

Hierbei stellt sich die Frage, in wie weit die Schwere des erlebten Automationsfehlers bezüglich etwa der Auswirkungen auf die eigene Gesundheit einen moderierenden Einfluss auf die Hemmung eines Cry Wolf Effektes hat.

Würde es hier keinen Zusammenhang geben, könnte eine Strategie für die Bildung eines Cry Wolf resistenten Rückmeldesystems darin bestehen, absichtlich „Automationsfehler“ ohne ernsthafte Konsequenzen zu simulieren. Die Erfahrung der Studie hat jedoch gezeigt, dass gerade dies zu einer verstärkten Aktivierung des extrinsischen Informationskanals führen kann – was zwar die Effektivität einer Rückmeldung nicht negativ beeinflusst – aber das Konzept der potenziellen Gefahr einer Situation negativ beeinflussen kann. So darf nicht der Eindruck entstehen, dass die Situationen, vor denen gewarnt wird, per se als nicht gefährlich wahrgenommen werden.

Auch im Vergleich zwischen Unsicherheit und einem idealen Alarm, führt das zweite Erleben eines Automationsfehlers zu einem deutlicheren Unterschied in der Reaktionszeit. Beim zweiten Automationsfehler reagieren die Versuchspersonen bei einem idealen Alarm besser als in der Unsicherheitsrückmeldung und ebenfalls besser als in der ersten Situation. Diese Verbesserung innerhalb des idealen Alarms, kann mit der gesteigerten allgemeinen Erwartung eines Automationsfehlers durch die Erfahrung des ersten Automationsfehlers erklärt werden.

## *10.2 Auswirkungen auf die allgemeine Erwartung von Automationsfehlern*

**Leithypothese 5 behauptet:** Wenn keine tatsächlichen Automationsfehler erlebt werden, führt die Rückmeldung möglicher Automationsfehler mittels Unsicherheit zu einem niedrigeren bzw. angemessenerem Vertrauen in die Automation, verglichen mit einer Bedingung ohne Rückmeldung, in der das Vertrauen sehr hoch ist.

Die Beobachtungen aus Studie 1 zeigen jedoch, dass allein die Rückmeldung von Unsicherheit, ohne das tatsächliche Erleben eines Automationsfehlers – nicht zur postulierten Anpassung des Vertrauens führt.

Mehrere Gründe können hierfür angeführt werden. Möglich ist z.B., dass gerade das Ausbleiben einer Konsequenz nach einer Rückmeldung zu einem stabilen Vertrauen in die Automation führt. Dies kann auch im Zusammenhang mit der bereits erwähnten Sekundärvarianz durch die hohe Gleichartigkeit der Rückmeldesituationen stehen. Findet die Rückmeldung in immer gleichen Situationen statt, in denen im Endeffekt nichts passiert, so ist es möglich, dass diese Situation als nicht mehr gefährlich wahrgenommen wird, da die Automation sie beherrschen kann. Das Vertrauen in die Automation steigt. In den Studien 2 und 3 wurde kein Vergleich mehr zwischen keiner Rückmeldung und Unsicherheitsrückmeldung vorgenommen, so dass eine Aussage unter kontrollierter Minimierung der möglichen Sekundärvarianz nicht vorgenommen werden kann.

**Leithypothese 6:** Ein Cry Wolf Effekt wirkt sich negativ auf das Vertrauen in die Automation aus. Wenn, bedingt durch einen Cry Wolf Effekt, etwa ein Automationsfehler schlecht beherrscht werden konnte, so ist das Vertrauen nach dieser Fahrt niedriger als nach einer Fahrt, in der durch keinen bzw. einen geringeren Cry Wolf Effekt der Automationsfehler besser beherrscht werden konnte. Folglich neigen Fahrer nach einer Fahrt in der Bedingung normaler Alarm eher zu Misstrust, als Fahrer nach einer Fahrt in der Bedingung Unsicherheitsrückmeldung. Diese Hypothese konnte in den durchgeführten Studien nur teilweise bestätigt werden.

In Studie 1 fand ein Vergleich zwischen Automationsfehler ohne vorherige Rückmeldung und Unsicherheitsrückmeldung statt. Hier zeigt sich, dass das Erleben eines Automationsfehlers – zu einem deutlichen Vertrauensverlust in die Automation führt, wenn vor diesem Fehler nicht vorher gewarnt wurde, dieser als vollkommen unerwartet passiert. In der Bedingung Unsicherheitsrückmeldung ist der Vertrauensverlust nach der gleichen Fahrt deutlich niedriger. Ebenfalls ist die abgefragte Tendenz die Automation nicht weiter nutzen zu wollen, nach einem

Automationsfehler ohne vorherige Rückmeldung deutlich niedriger als in der Bedingung Unsicherheit.

In Studie 2 wurde der Vergleich zwischen einem normalen Alarm und der Unsicherheitsrückmeldung durchgeführt – jedoch kam es hier, wahrscheinlich aufgrund der erwähnten Sekundärvarianz, kaum zu Unterschieden in der Kontrollierbarkeit der Automationsfehler und somit bei beiden Rückmeldearten zu einem niedrigen Gesamtvertrauen.

Auch wenn für die geringe Erwartung des Automationsfehlers nicht direkt ein Cry Wolf Effekt verantwortlich ist, sondern wohl eher die gering wahrgenommene Gefahr der Situation – so sind die daraus resultierenden Konsequenzen vergleichbar. Eine geringe Erwartung eines Automationsfehlers macht diesen weniger gut kontrollierbar, entsprechend sinkt das Gesamtvertrauen in die Automation rapide.

In Studie 3 wurde besagte Sekundärvarianz zwar kontrolliert, das Vertrauen in die Automation aber aus Gründen der Versuchsdurchführung erst nach dem zweiten erlebten Automationsfehler gemessen. Durch das vorherige Erleben eines Automationsfehlers wurde die Anfälligkeit für einen Cry Wolf Effekt auch in der Bedingung normaler Alarm deutlich gesenkt. Entsprechend vergleichbar war die Kontrollierbarkeitsreaktion des zweiten Automationsfehlers zwischen den Bedingungen normaler Alarm und Unsicherheit. In beiden Bedingungen war das Vertrauen in die Automation entsprechend hoch.

Summiert bestätigt sich die Hypothese, dass eine geringe Erwartung eines Automationsfehlers, z.B. verursacht durch einen Cry Wolf Effekt, oder allgemeiner durch Overtrust, zu einer schlechten Kontrollierbarkeit des Fehlers führt. Die nicht Vorhersehbarkeit von Fehlern führt zu einem negativen Gesamtvertrauen in die Automation. Wird ein Fehler gar nicht vorhergesehen, wenn z.B. gar nicht vor einem Fehler gewarnt wird (Miss Situation in Studie 3) ist das Vertrauen in die Automation deutlich negativ beeinflusst, und zwar unabhängig von der sonst erlebten Rückmeldungsart. Fehler ohne vorherige Warnung sollten daher nicht nur aus Gründen der geringeren Kontrollierbarkeit, sondern auch aufgrund der deutlich negativen Auswirkung auf das Systemvertrauen vermieden werden.

Es konnte in dieser Arbeit nicht herausgestellt werden, dass sich eine Unsicherheitsrückmeldung aufgrund der geringeren Cry Wolf Anfälligkeit hemmend auf die Entstehung von Misstrust nach einem erlebten Automationsfehler auswirkt. Dies ist aber eher auf methodische Ursachen zurückzuführen und sollte in zukünftigen Untersuchungen näher betrachtet werden.

**Leithypothese 7:** Das Vorhandensein eines Rückmeldesystems das vor Automationsfehlern warnt, darf nicht zu einer schlechteren Kontrollierbarkeit eines Automationsfehlers führen, wenn vor diesem keine Rückmeldung erfolgt verglichen

mit einer Bedingung ohne Rückmeldung. Die Compliance, d.h. die Erwartung, dass ein Fehler nur auftritt, wenn vorher auch eine Rückmeldung erfolgte, darf nicht höher sein als die allgemeine Erwartung eines Automationsfehlers ohne Rückmeldesystem. Dieser Hypothese wurde in exploratorischer, nicht streng kontrollierter Weise nachgegangen. Die Versuchspersonen erlebten in Studie 3 eine Miss Situation zum Abschluss der Fahrt, also ein Automationsfehler, vor dem keine Rückmeldung erfolgt, obwohl aber ein Alarm/Warnsystem vorhanden ist. Eine Vergleichsbedingung eines Automationsfehlers vor dem aber generell keine Rückmeldung erfolgt war in Studie 1 vorhanden. Trotz der ähnlichen Situation ist eine eindeutige Vergleichbarkeit daher nicht gegeben. Die Kontrollierbarkeit der Miss Situation war insgesamt schlechter als in der Situation des zweiten Automationsfehlers in Studie 3 und zwar in gleicher Art für alle Rückmeldungsbedingungen. Die Reaktionszeit war aber nicht schlechter als in der Bedingung ohne Rückmeldesystem in Studie 1.

Unter Vorbehalt kann daher angenommen werden, dass das Vorhandensein eines Systems, das auf mögliche Automationsfehler hinweist, nicht zu allgemein geringerer Erwartung eines Automationsfehlers als in einer Bedingung ganz ohne ein solches Rückmeldesystem führt. Jedoch ist es wahrscheinlich, dass in beiden Fällen ein Automationsfehler gar nicht erwartet wurde – eine noch geringere Erwartung als keine Erwartung ist nicht möglich.

## **11 Schlussfolgerung**

Zusammengefasst kann gefolgert werden, dass es einen Unterschied in der Verarbeitung bzw. Wirkweise einer Unsicherheitsrückmeldung und einem normalen Alarm gibt. Insgesamt kann auch festgestellt werden, dass eine Unsicherheitsrückmeldung effektiver hinsichtlich der Übernahmefähigkeit bei Automationsfehlern ist als ein konventioneller Alarm, wenn man den Einfluss des Cry Wolf Effektes mit einbezieht und auf jeden Fall effektiver als keine Rückmeldung. Dieser Vorteil gegenüber konventionellen Alarmen ist aber nicht absolut zu sehen, sondern hängt von Rahmenbedingungen ab.

- Sie hängt davon ab, ob echte Alarmer tatsächlich erlebt wurden oder nicht.
- Sie hängt von der Art des erlebten Automationsfehlers und der Ähnlichkeit der Situationen mit Automationsfehler ab.

Auch eine Unsicherheitsrückmeldung ist nicht immun gegenüber der Entstehung eines Cry Wolf Effektes. Sie weist sicherlich eine höhere Resistenz auf als ein normaler Alarm, jedoch scheint das Ausbleiben eines Automationsfehlers nach einer Rückmeldung nicht, wie vermutet, gar keine negative Evidenz darzustellen. Die negative Evidenz wird wohl eher geringer gewichtet als bei einem normalen Alarm.

Die Studien konnten insgesamt deutliche Hinweise liefern, dass die Theorie der Evidenzevaluation nach White (z.B. 2003) zur Erklärung unterschiedlicher Effekte der Rückmeldungsgestaltung auf die Entstehung eines Cry Wolf Effektes herangezogen werden kann. Dies liefert Basis für zukünftige theoriegeleitete Gestaltung von Rückmeldungsstrategien.

Die Beobachtungen der Studien stützen die Hypothese, dass Rückmeldungen und das Erleben bzw. das nicht Erleben zur Bildung und Veränderung von Zusammenhangshypothesen führt. Die Hypothese konnte gestützt werden, dass im Allgemeinen durch die Art der Rückmeldungsgestaltung ein Einfluss auf die Art der Zusammenhangshypothese genommen werden kann. Auch, dass das Ausbleiben einer Konsequenz nach einer Rückmeldung in Abhängigkeit der Rückmeldungsgestaltung, eine anders geartete Evidenz darzustellen scheint. Dabei kann nicht einfach zwischen positiver und negativer Evidenz unterschieden werden, sondern es muss hier auch eine Art Gewichtung der Evidenz in Erwägung gezogen werden.

Im Hinblick auf die generelle Erwartung eines Automationsfehlers, kann aufgrund der Studien die Hypothese nicht gestützt werden, dass es durch die Rückmeldung von Automationsunsicherheit zu einer Bildung eines allgemeinen adäquaten Vertrauens in die Automation kommt –welches implizit die Erwartung eines Automationsfehlers einschließt, auch wenn ein solcher nicht erlebt wurde. Die größte Auswirkung auf die allgemeine Erwartung von Automationsfehlern hat die Erfahrung tatsächlicher Automationsfehler!

Aufgrund von methodischen Implikationen konnte nicht belegt werden, dass die Rückmeldung von Unsicherheit dem Entstehen von Misstrust in die Automation vorbeugen kann. Es konnten aber allgemeine Hinweise gefunden werden, dass sich eine schlechtere Kontrollierbarkeit von Automationsfehlern negativer auf das Gesamtvertrauen auswirkt als eine gute Kontrollierbarkeit. Zur Vorbeugung der Entstehung von Misstrust in die Automation sollte daher vor allem vermieden werden, dass Automationsfehler, insbesondere mit möglichen schwerwiegenden Konsequenzen, vollkommen unerwartet passieren. Möglich ist hier ein Zusammenhang in der Art, dass das Vertrauen in die Automation umso negativer beeinflusst ist, desto mehr ein Fahrer bzw. Automationsnutzer durch einen Automationsfehler überrascht wird.

Insgesamt wird gefolgert, dass eine situative Anpassung der Erwartung der Leistungsfähigkeit einer Automation durch eine Rückmeldung welche auch noch als eine Art Begründung für die geringere Leistungsfähigkeit angesehen werden kann – einen wesentlichen Baustein für eine tatsächliche kooperative Fahrzeugführung darstellen kann. Insbesondere vor dem Hintergrund – dass die Fahrzeuge vermeintlich intelligenter und auch durch die Fahrer so wahrgenommen werden –

entsteht eine Neigung der Fahrer die technischen Systeme zu vermenschlichen. Technischen Systemen werden menschliche Eigenschaften zugeschrieben und entsprechend werden die Systeme auch behandelt. (Nass & Moon, 2000, Reeves & Nass, 1996)

Insbesondere vor diesem Hintergrund scheint es angebracht – eine Rückmeldung wie die Unsicherheitsrückmeldung zu nutzen – welche, ähnlich der Mensch-Mensch Kommunikation, mitteilt wenn ein Kooperationspartner unsicher ist, in einer bestimmten Situation eine ausreichende Hilfe darzustellen. Der andere Partner kann sich dann entsprechend drauf einstellen, eventuell größere Teile der Aufgabe übernehmen zu müssen. Der auf diese Weise transparente Einblick in die Leistungsfähigkeit schafft somit größeres Vertrauen in ein System – als weniger transparente und wahrscheinlich auch weniger verständliche Kontrollübernahme-Aufforderungen.

## **12 Ausblick**

Die Erfahrungen der durchgeführten Studien haben teils ein deutliches Potenzial für weitere Forschung aufgezeigt.

- Was sind z.B. die Merkmale einer Situation, welche eine Situation ähnlich bzw. gleich erscheinen lassen und somit einen extrinsischen Informationsfluss fördern? Bzw. bei welcher Unterschiedlichkeit der Situationen wird eher ein intrinsischer Informationsfluss gefördert, so dass eine Attribution auf das Alarmsystem erfolgt? Hier können z.B. Konzepte des intuitiven Urteilens, wie z.B. die Repräsentativitätsheuristik „sieht aus wie..., muss auch wie... sein“, herangezogen werden (z.B. Tversky & Kahnemann, 1973).
- Auswirkungen unterschiedlicher Alarmstrategien auf die Strategie der Aufmerksamkeitsverteilung bei zwei konkurrierenden Aufgaben (Hauptaufgabe und Nebenaufgabe).
- Erforschung weiterer Methoden zu Messung situativen Trusts bzw. der situativen Erwartung von Automationsfehlern oder anderer Ereignisse.
- Erforschung weiterer Möglichkeiten einen Cry Wolf Effekt zu hemmen z.B. durch die Simulation von Automationsfehlern während einer Fahrt.
- Systematische Erforschung der Interaktion aus Fehlalarmhäufigkeit und Zeit in der diese erlebt werden.
- Systematische Erforschung der Bedeutung der Gewichtung von Evidenzen bezüglich der Entstehung eines Cry Wolf Effektes und der Effektivität einer Rückmeldung.
- Weitere Untersuchung der Auswirkungen der Bereitstellung einer Begründung für Fehler – auf das Vertrauen in die Automation.

## 13 Referenzen

- Adams, M.J., Tenney, Y.J. & Pew, R.W. (1995) Situation awareness and the cognitive management of complex systems. *Human Factors*, 37(1), 85-104
- Antifakos, S., Kern, N., Schiele, B. & Schaninger, A. (2005) Towards improving trust in context-aware systems by displaying system confidence. Published 2005 in *Mobile HCI*
- Bagheri, N. & Jamieson, G.A. (2004) The Impact of Context-related Reliability on Automation Failure Detection and Scanning Behaviour *IEEE International Conference on Systems, Man and Cybernetics*. Vol. 1. pp.212-17.
- Bahner, J. (2008) Übersteigertes Vertrauen in Automation: Der Einfluss von Fehlererfahrungen auf Complacency und Automation Bias. Dissertation an der Fakultät für Verkehrs- und Maschinensysteme der TU Berlin
- Banbury, S., Selcon, S., Endsley, M., Gorton, T. & Tatlock, . (1998). Being Certain About Uncertainty: How the Representation of System Reliability Affects Pilot Decision Making. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 42. 36-39.
- Barber, B. (1983). *The Logic and Limits of Trust*. Rutgers University Press, New Brunswick, NJ.
- Baumann, M. & Krems, J. F. (2009). A Comprehension Based Cognitive Model of Situation Awareness. In V. D. Duffy (ed.), *Digital Human Modeling*, Vol. 5620 (pp. 192--201). Springer, New York.
- Baumann, M., & Krems, J. (2005). Situation awareness and driving: a cognitive model of memory update-processes. In L. Macchi, Chr. Re & P.C. Cacciabue (Eds.). *Proceedings of the International Workshop on Modelling Driver Behaviour in Automotive Environments* (pp. 157-163). Luxembourg: Publication of the European Communities.
- Benesch, M. & Raab-Steiner, E. (2013). *Klinische Studien lesen und verstehen*. Wien: Facultas.
- Bisantz, A.M., Marsiglio, S.S. & Munch, J. (2005) Displaying uncertainty: investigating the effects of display format and specificity. *Human Factors*, 47(4), 777-796.
- Bliss, J.P. & Acton, S.A. (2003) Alarm mistrust in automobiles: how collision alarm reliability affects driving. *Applied Ergonomics* 34 (2003) 499–509
- Bliss, J.P. & Dunn, M.C. (2000). Behavioural implications of alarm mistrust as a function of task workload, *Ergonomics*, 43 (9), 1283-1300.



- Bliss, J, Dunn, M. & Fuller, B.S. (1995) Reversal of the cry-wolf effect: an investigation of two methods to increase alarm response rates. *Perceptual and Motor Skills*, 80 pp.1231-1242
- Bliss, J., Gilson, R., & Deaton, J. (1995). Human probability matching behaviour in response to alarms of varying reliability. *Ergonomics*, 38, 2300–2313.
- Brandenburg, S., & Skottke, E. M. (2014). Switching from manual to automated driving and reverse: Are drivers behaving more risky after highly automated driving? In *Intelligent Transportation Systems (ITSC)*, 2014 IEEE 17th International Conference on (pp. 2978- 2983). IEEE.
- Bresnitz, S. (1983). *Cry-wolf: The psychology of false alarms*. Hillsdale, NJ: Lawrence Erlbaum.
- Bustamante, E.A. (2008) Implementing Likelihood Alarm Technology in Integrated Aviation Displays for Enhancing Decision-Making: A Two- Stage Signal Detection Modeling Approach *International Journal of Applied Aviation Studies*, 8 (2)
- Cafarelli, D.A. & Hansmann, R.J. (1998) Effects of False Alarm Rate on Pilots use and Trust of Automation under conditions of simulated high risk. Retrieved from: <http://dspace.mit.edu/bitstream/handle/1721.1/59842/Cafarelli%20Paper-ICAT-1998-04.pdf?sequence=1>
- Christoffersen, K, & Woods, D.D. (2002) How to Make Automated Systems Team Players. *Advances in Human Performance and Cognitive Engineering Research*,2, 1–12
- Clark, R.M., Ingebritsen, A.M. & Bustamante, E.A. (2010) Differential Effects of Likelihood Alarm Technology and False-Alarm vs. Miss-Prone Automation on Decision-Making Accuracy and Bias. *Proceedings of the Human Factors and Ergonomics Society 54th annual meeting – 2010*
- Damböck, D., Farid, M., Tönert, L. & Bengler K. (2012) Übernahmezeiten beim hochautomatisierten Autofahren. 5. Tagung Fahrerassistenz 2012. München. Germany.
- Darms, M. (2007) Eine Basis-Systemarchitektur zur Sensordatenfusion von Umfeldsensoren für Fahrerassistenzsysteme. Dissertation am Fachbereich Maschinenbau an der Technischen Universität Darmstadt
- De Waard, D., van der Hulst, M., Hoedemaeker, M. & Brookhuis, K.A. (1999) Driver Behavior in an Emergency Situation in the Automated Highway System, *Transportation Human Factors*, 1:1, 67-82,
- Dixon, S.R., Wickens, C.D. & McCarley, J.S. (2006) How do Automation False Alarms and Misses Affect Operator Compliance and Reliance? *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting—2006*

- Dixon, S.R., Wickens, C.D. & McCarley, J.S.(2007) On the independence of compliance and reliance: are automation false alarms worse than misses? *Human Factors*. 49 (4), 564-72.
- Dominguez, C. (1994). Can SA [Situation Awareness] be defined? In M. Vidulich (Ed.). *Situation Awareness: Papers and Annotated Bibliography*. Armstrong Laboratory
- Duden, Redewendungen. Wörterbuch der deutschen Idiomatik. In: Der Duden in zwölf Bänden. 2., neu bearbeitete und aktualisierte Auflage. Band 11, Dudenverlag, Mannheim/Leipzig/Wien/Zürich 2002
- Duschl, M., Just, M. & Niedermaier, B. (2011) Development and evaluation of a visualization for displaying uncertain data. User Evaluation of uncertainty visualizations in context of a deceleration assistance system. *VDI-Berichte* Band:2134 Seite:169-181 VDI-Verlag Düsseldorf
- Endsley, M.R. (1995): Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors Journal* 37(1), 32-64
- Endsley M. R. & Kiris, E.O. (1995) The out-of-the-loop performance problem and level of control in automation. *Human factors*. 37(2), 381-394.
- Finger, R., & Bisantz, A. M. (2002). Utilizing graphical formats to convey uncertainty in a decision-making task. *Theoretical Issues in Ergonomics Science*, 3(1), 1-25.
- Flemisch, F. Heesen, M., Hesse, T., Kelsch, J., Schieben, A. & Beller, J. (2012) Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative control situations *Cognition, Technology & Work* 14 (1), 3-18
- Gasser, T.M., Arzt, C., Ayoubi, M., Bartels, A., Eier, J., Flemisch, F., Häcker, D., Hesse, T., Huber, W., Lotz, C., Maurer, M., Ruth-Schumacher, S., Schwarz, J., Vogt, W. (2012). Projektgruppe "Rechtsfolgen zunehmender Fahrzeugautomatisierung" BASt-Bericht F 83.
- Getty, D.J., Swets, A., Pickett, R.M., & Gonthier, D. (1995). System Operator Response to Warning of Danger: A Laboratory Investigation of the Effects of the Predictive Value of a Warning on Human Response Time. *Journal of Experimental Psychology: Applied*, 1, 19-33.
- Girden, E. (1992). *ANOVA: Repeated measures*. Newbury Park, CA: Sage.
- Gold, C. & Bengler, K. (2014) Taking over control from highly automated vehicles. In N. Stanton, S. Landry, G. Di Bucchianico & A. Valicelli (Hrsg.) *Advances in Human Aspects of Transportation: Part II*
- Gomes; L. (2014, August 28). Hidden Obstacles for Google's Self-Driving Cars [Article on MIT technology review - new]. Retrieved from

- <http://www.technologyreview.com/news/530276/hidden-obstacles-for-googles-self-driving-cars/>
- Green, M. (2000). "How Long Does It Take to Stop?" Methodological Analysis of Driver Perception-Brake Times. *Transportation Human Factors* 2(3)
- Green, D.M., Swets J.A. (1966) *Signal Detection Theory and Psychophysics*. New York: Wiley
- Grigoryan, G. & Rheingans, P. (2002) "Probabilistic surfaces: point based primitives to show surface uncertainty." *IEEE Visualization*: 147-153.
- Hagmeyer, York & Waldmann, Michael R. in Funke, Joachim (Hrsg) *Enzyklopädie der Psychologie / Themenbereich C: Theorie und Forschung / Kognition / Denken und Problemlösen Gebundenes Buch – 1. August 2006 Hogrefe*
- Helldin, T., Falkman, G., Riverio, M. & Davidson, S. (2013) Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. Conference: Conference: Proc. 5th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 13), At Eindhoven, The Netherlands
- Hoc, M.J., Young, M.S. & Blosseville, J.M. (2009) Cooperation between drivers and automation: Implications for safety Theoretical Issues in Ergonomics Science
- Johnson-Laird, P (1983) *Mental Models*, Harvard University Press
- Kassner, A. & Baumann, M. (2011) Situationsbewusstsein bei unterschiedlichen Automationsgraden. In: Beiträge zur 53. Tagung experimentell arbeitender Psychologen, Seite 86. Pabst Science Publishers
- Kraiss, K.-F. (1994). 99% Langeweile und 1% panische Angst“ –über die Schwierigkeiten beim Umgang mit hochautomatisierten Systemen. In M. Kerner (Ed.), *Technik und Angst. Zur Zukunft der industriellen Zivilisation*. Aachen: Verlag der Augustinus Bhg.
- Lee, J.D. & See, K.A. (2004) Trust in Automation, Designing for Appropriate Reliance. *Human Factors*. (46(1), 50-80
- Markoff, J. (2014, May 27). Google's Next Phase in Driverless Cars: No Steering Wheel or Brake Pedals. *New York Times*. Retrieved from [http://www.nytimes.com/2014/05/28/technology/googles-next-phase-in-driverless-cars-no-brakes-or-steering-wheel.html?\\_r=0](http://www.nytimes.com/2014/05/28/technology/googles-next-phase-in-driverless-cars-no-brakes-or-steering-wheel.html?_r=0)
- Macmillan, N. A. (2002) Signal detection theory. In: Pashler, H. (Hrsg.), *Stevens' Handbook of Experimental Psychology*, Band 1, Kapitel 2, Seiten 43–91. Wiley, New York, 3. Auflage.

- Masalonis, A.J. & Parasuraman, R. (1999). Trust as a Construct for Evaluation of Automated Aids: Past and Future Theory and Research. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 43. 184-187.
- Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. Paper pre-sented at the Annual Spring Conference of the GfA/17th Annual Conference of International-Society-for-Occupational-Ergonomics-and-Safety (ISOES).
- Master, R., Gramopadhye, A., Melloy, B., Bingham, J., & Jiang, X. (2000). A Questionnaire for Measuring Trust in Hybrid Inspection Systems. *Proceedings of the Industrial Engineering Research Conference, Dallas TX*,
- McGuirl, J. & Sarter, N. (2006). Supporting Trust Calibration and the Effective Use of Decision Aids by Presenting Dynamic System Confidence Information. *Human factors*. 48. 656-65
- Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., & Bauer, W. (2015). Take-over requests for automated driving. *Procedia Manufacturing*, 3, 2867-2873.
- Merat, Natasha & Jamson, A. (2009). How do drivers behave in a highly automated car. Paper downloaded from Researchgate.
- Merat, N., Jamson, A. H., Lai, F. C., Daly, M., & Carsten, O. M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation research part F: traffic psychology and behaviour*, 27, 274-282.
- Meyer, J. (2001). Effects of warning validity and proximity on responses to warnings. *Human Factors*, 43 (4), 563-572.
- Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors*, 46 (2), 196-204.
- Muir, B.M., 1994. Trust in automation: part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics* 37, 1905–1922.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81-103.
- Neukum, A., Lübbecke, T., Krüger, H.P., Mayser, C. & Steinle, J. (2008). ACC-Stop&Go: Fahrerverhalten an funktionalen Systemgrenzen. In M.Maurer & C.Stiller (Hrsg.), 5.Workshop Fahrerassistenzsysteme – FAS 2008 (S.141-150). Karlsruhe: fmrt.
- NHTSA, National Highway Traffic Safety Administration (2013). Preliminary Statement of Policy Concerning Automated Vehicles.  
[www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated\\_Vehicles\\_Policy.pdf](http://www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated_Vehicles_Policy.pdf)  
 [Abgerufen am 26.06.2018)

- Niederée, U & Vollrath, M. (2009). Systemausfälle bei Längsführungsassistenten- Sind bessere Systeme schlechter? Vortrag bei der 8. Berliner Werkstatt Mensch-Maschine-Systeme „Der Mensch im Mittelpunkt technischer Systeme“, Berlin.
- Parasuraman, R. , Hancock, P. A. and Olofinboba, O.(1997) 'Alarm effectiveness in driver-centred collision-warning systems', *Ergonomics*, 40: 3, 390 — 399
- Parasuraman, R., Molloy, R. & Singh, I. L. (1993) Performance Consequences of Automation-Induced 'Complacency' *The International Journal of Aviation Psychology* Vol 3, No 1
- Parasuraman, R. & Riley, V. (1997) Humans and Automation, Use, Misuse, Disuse, Abuse. *Human Factors*, 39(2), 230-253
- Petermann-Stock, I., Hackenberg, L., Muhr, T., & Mergl, C. (2013). Wie lange braucht der Fahrer--eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt. 6. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren.
- Rasmussen J. (1986) *Information processing and human-machine interaction*, Amsterdam: Elsevier North Holland.
- Rauch, N., Gradenegger, B. & Krüger, H.P. (2009). Darf ich oder darf ich nicht? Situationsbewusstsein im Umgang mit Nebenaufgaben während der Fahrt *Zeitschrift für Arbeits. Wissenschaft.* (63)
- Rempel, J.K., Holmes, J.G., Zanna, M.P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology* 49, 95–122.
- Reeves, B. & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Pla.* Bibliovault OAI Repository, the University of Chicago Press.
- Roulston, M.S. & Smith, L.A. (2004) The Boy Who Cried Wolf Revisited: The Impact of False Alarm Intolerance on Cost–Loss Scenarios. *Weather & Forecasting*. Apr 2004, Vol. 19 Issue 2, p391-397
- Rudin-Brown, C.M., Parker, H.A. (1995) Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies *Transportation Research Part F: Traffic Psychology and Behaviour* Volume 7, Issue 2, March 2004, Pages 59-76
- SAE, Society of Automotive Engineers (2014). Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. *Surface Vehicle Information Report J3016*, 1-12. [http://standards.sae.org/j3016\\_201401/](http://standards.sae.org/j3016_201401/) Stand Juni 2018
- Sarter, N.B. & Woods, D.D. (1995) How in the world did we ever get into that mode?: Mode error and awareness in supervisory control. *Human Factors*, 37. pp. 5-19

- Sarter, N.B., Woods, D.D. & Billings, C.E. (1997) Automation Surprises. In: G. Salvendy (Ed.) Handbook of Human Factors & Ergonomics, second edition, Wiley.
- Satchell, P. (1993), Cockpit Monitoring and Alerting Systems (Ashgate, Aldershot).
- Seppelt, B.D. (2009) Supporting Operator Reliance On Automation Through Continuous Feedback. Doctoral Thesis, University of Iowa.
- Seppelt, D. & Lee, J.D. (2007) Making adaptive cruise control (ACC) limits visible. International Journal of Human-Computer Studies vol. 65. 192–205
- Skjerve, A.B.M. (2002). The Halden Co-operation Scale. Human-Automation Co-operation in Control Room Settings. HWR-685, OECD Halden Reactor Project, Halden, Norway.
- Smith, K., & A., P. (1995). Situation Awareness Is Adaptive, Externally Directed Consciousness. Human Factors, 37(1), 137–148.
- Sorkin, R.D., Kantowitz, B.H. & Kantowitz, S.C. (1988). Likelihood Alarm Displays. Human Factors, 30 (4), 445-459
- Toffetti, A., Wilschut, E. S., Martens, M. H., Schieben, A., Rambaldini, A., Merat, N., & Flemisch, F. (2009). CityMobil. Transportation Research Record: Journal of the Transportation Research Board, 2110(1), 1-8.
- Totzke, I., Volks, M., Naujoks, F. & Krüger, H.P. (2013). Unzuverlässige Informationen über die Positionierung eines Stauendes: Wie wirken sich falsche Distanzangaben auf das Fahrerverhalten aus? AAET 2013, Braunschweig
- Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 42, 207–232.
- Vogelpohl, T., Vollrath, M., Kühn, M., Hummel, T. & Gehlert, T. (2016) Übergabe von hochautomatisiertem Fahren zu manueller Steuerung Teil 2: Müdigkeit und lange Fahrtdauer als Einflussfaktoren auf die Sicherheit nach einer Übernahmeaufforderung- Berlin: GDV, 2016
- Vogelpohl, T., Vollrath, M. Kühn, M., Hummel, T. & Gehlert, T (2016) Übergabe von hochautomatisiertem Fahren zu manueller Steuerung Teil 1: Review der Literatur und Studie zu Übernahmezeiten- Berlin: GDV, 2016
- Wiczorek, R. & Manzey, D. (2011) Verhaltenswirksamkeit von Likelihood Alarmsystemen. Fortschritt Berichte VDI Reihe 22 Nr. 33
- Wiczorek, R. (2011). Entwicklung und Evaluation eines mehrdimensionalen Fragebogens zur Messung von Vertrauen in technische Systeme. In S. Schmid, M. Elepfandt, J. Adenauer, & A. Lichtenstein (Eds.), Reflexionen und Visionen der Mensch-Maschine-Interaktion –Aus der Vergangenheit lernen, Zukunft gestalten (pp. 198–199). 9. Berliner Werkstatt Mensch-Maschine-Systeme, Berlin: VDI.

- White, P. A. (2000). Causal judgment from contingency information: Relation between subjective reports and individual tendencies in judgment. *Memory & Cognition*, 28, 415- 426.
- White, P. A. (2002). Causal attribution from covariation information: The evidential evaluation model. *European Journal of Social Psychology*, 32, 667-684.
- White, P. A. (2003). Making causal judgments from the proportion of confirming instances: The pCI rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 710-727.
- Wickens, C.D. (1996). Situation awareness: Impact of automation and display technology. AGARD Conference Proceedings 575: Situation Awareness: Limitations and Enhancement in the Aviation Environment (pp. K2-1 - K2-13). Neuilly-Sur-Seine, France: Advisory Group for Aerospace Research & Development.
- Wickens, C. D. & Dixon, S.R. (2007) The benefits of imperfect diagnostic automation: a synthesis of the literature. *Theoretical Issues in Ergonomic Science* Vol8, No. 3, May-June 2007, 201-212
- Wiener, E. L. (1988), Cockpit automation, in E. L. Wiener and D. C. Nagel (eds), *Human Factors in Aviation* (Academic Press, San Diego, CA), 433-461.
- Wulf, F., Rimini-Döring, M., Arnon, M., & Gauterin, F. (2014). Recommendations Supporting Situation Awareness in Partially Automated Driver Assistance Systems. *IEEE Transactions on Intelligent Transportation Systems*, Vol.PP, no.99, pp.